Iterative methods for heterogeneous media

Lechner, Patrick O.

Award date:
2006

Awarding institution:
University of Bath

Iterative Methods for Heterogeneous Media

submitted by

Patrick O. Lechner

for the degree of Doctor of Philosophy

of the

University of Bath

Department of Mathematical Sciences

February 2006

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

To my parents

Summary

Heterogeneous media play a key role in the analysis and computation of groundwater flows. Due to their underlying locally strongly varying permeability fields the discretisation of these problems often leads to highly ill-conditioned large sparse systems. These are hard to solve iteratively and it is essential to find good preconditioners. Domain decomposition methods with linear interpolation however often perform poorly when the variation of the coefficients inside the subdomains is large.

For the modeling of these problems lognormal random fields of Ornstein-Uhlenbeck type are often used to capture the heterogeneity of the medium and to deal with uncertainties between the finite number of measure points of the fields. In this thesis we therefore first of all analyse these fields and derive bounds of the probability distributions of maxima of global and local field ratios using Poisson clumping heuristics. We then prove that these ratios characterise the dependence of upper bounds of the condition numbers of the unpreconditioned and linear-interpolation based additive Schwarz preconditioned systems on the permeability.

To reduce this dependence we introduce a new coarsening operator based on multiscale finite elements (MsFE). The significance of the choice of the boundary conditions for the construction of the MsFE-basis functions to achieve good convergence rates is discussed in detail. By considering conjugate gradient iteration numbers and computation times for 1D, 2D and 3D domains and two-media problems as well as realisations of random permeability fields we show that the new coarsening operator outperforms linear interpolation in most cases. We prove that the new preconditioner is in fact the exact inverse of the stiffness matrix for 1D problems and that the same is true for a skeleton-based extension of it for higher-dimensional problems. A complexity analysis of the different methods, that are considered in this thesis, finally try to assist in the choice of the right preconditioner for a given problem.

Acknowledgements

It is a great pleasure to thank the people without whom this thesis never would have been possible.

It is difficult to overstate my gratitude to my Ph.D. supervisor, Ivan Graham. His enthusiasm, his encouragement, whenever I struggled, and all his support mean the world to me. He gave me ideas and advice and most importantly he made my Ph.D. a simply fantastic and unforgettable time. I would have been lost without him.

I would like to thank the people, whose expertise and input helped me a lot over the last few years. First and foremost, I wish to thank Robert Scheichl, who gave me so many new ideas, who was there for me, whenever I needed his help, and who became a true friend. I am grateful to Mathew Penrose and Merrilee Hurn for their support in the analysis of random fields and to Valery Smyshlyaev, Eero Vainikko, Adrian Hill, Alastair Spence, Bill Morton and many other members of the Department of Mathematics for all the good advice they gave me and all the nice chats we had. I will miss this a lot.

I am also indebted to many colleagues from outside the University of Bath for fruitful discussions and useful references, especially to Susanne Brenner, Mary Wheeler, Petr Plechac, Yalchin Efendiev, Wolfgang Hackbusch, Mario Bebendorf, Tom Hou, David Keyes, Frederic Nataf, Endre Suli and Andrew Cliffe. A great thank you also goes to Benjamin Kirk for his support with Libmesh.

My Ph.D. would however have been only half as much fun without my brilliant office mates, past and present, Andy, Andre, Bradley, Dave, Laura, Melina, Petra, Stefano and Simone (who was the first person I met in Bath, then survived four years in the same office with me and despite spending such a long time in my company still became a great friend). Thanks to all of you for putting up with me and for many wonderful memories! I would like to thank Bob, Mary, Jill, Dawn, Dena and the rest of the computing support and administration for always being there for me.

I am very grateful to the Department of Mathematics for sponsoring my Ph.D.. The University of Bath gave me so much help and support during the last few years and showed me what great fun research can be. It gave me so much more during my MSc and Ph.D. than I could ever have expected, when I came here four years ago and I will never ever forget this fantastic hospitality. Bath will always be in my heart!

Special thanks go to my other friends, that I made here in Bath over the last four years. Friends make life fun and I am a lucky guy to have lived with Pete, James, Paul, Jo, Matt and Nicola during the last three years. Next to them I would like to thank in particular Michelle, Richard, Nathan, Stephanie, Carl-Friedrich, Sally, Helen, Gundula and Adam for many wonderful moments. Thanks for making sure that I did not go completely insane while writing this thesis!

Friendship over long distances is not always easy and I am therefore especially grateful that Alexander has been such a great friend in Dellmensingen, resp. Biberach and Bonn and has still remained of fantastic support and encouragement despite living 700 kilometers away now. Whenever I need help, he is there for me. Danke Alex!

I wish to thank my entire extended family for providing a loving environment for me. Special thanks go to my sister, Isabelle, who I miss a lot.

Lastly and most importantly, I would like to thank my parents, Hedi and Herwig Lechner. They have always believed in me, even when I did not, they support me and give me strength and love. I would be nothing without them. To my parents I dedicate this thesis.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

" *Water is not a luxury. Actually, lack of water is more costly.*"

— Michael Okema, political scientist, Dar es Salaam, Tanzania, 2004 —

1.1 The subject of the thesis

Water is the basis of all life and groundwater the biggest and most important occurrence of fresh water on earth. As this resource is very scarce in many countries and as conflicts about water are sadly starting to become reality, it seems the duty of scientists of various disciplines to learn more about it.
Mathematicians often study processes in continuum physics and engineering with the help of models based on partial differential equations (PDEs). The flow of water underground can then for example be described by the first order system

$$\begin{aligned} \mu\alpha^{-1}\mathbf{v} + \boldsymbol{\nabla} u &= \mathbf{0}, \\ \boldsymbol{\nabla}.\mathbf{v} &= 0, \end{aligned} \tag{1.1}$$

which has to be solved for a (vector) *velocity* $\mathbf{v}$ and a (scalar) *pressure* u on an (open) domain $\Omega \subset \mathbb{R}^d$. The scalar coefficients μ and $\alpha(\mathbf{x})$, $\mathbf{x} \in \Omega$, are called the *dynamic viscosity* of the fluid and, respectively, the *permeability* of the underlying porous medium. (In many models α will in fact be a full tensor, but we do not consider that case here.) System (1.1) has to be solved subject to boundary conditions

$$u = u_0 \text{ on } \Gamma_D \quad \text{and} \quad \mathbf{v}.\boldsymbol{\nu} = 0 \text{ on } \Gamma_N, \tag{1.2}$$

where Γ_D and Γ_N form a non-overlapping decomposition of Γ, the boundary of Ω, and where $\boldsymbol{\nu}(\mathbf{x})$ denotes the outward unit normal from Ω at $\mathbf{x} \in \Gamma_N$. The first of these two boundary conditions is often also referred to as a *Dirichlet*, the second as a *Neumann boundary condition.*

System (1.1, 1.2) can be discretised in many different ways. A good discretisation however should have a unique solution, which is "close" to the exact solution of (1.1, 1.2) in terms of local and global errors and which should display some physical aspects of the underlying problem. For undergroundwater flow problems, this often means that mass should be conserved locally, i.e. we assume a balance of the amount of water that flows in and out of a small fixed region. This can be achieved for example by searching solution vectors $\mathbf{U}$ and $\mathbf{V}$ in two different approximation spaces for u and $\mathbf{v}$, in a so called mixed finite element method, which leads to the discrete system

$$\begin{bmatrix} W(\alpha) & B \\ B^T & 0 \end{bmatrix} \begin{pmatrix} \mathbf{V} \\ \mathbf{U} \end{pmatrix} = \begin{pmatrix} \mathbf{G} \\ \mathbf{0} \end{pmatrix}. \tag{1.3}$$

Here $W(\alpha)$ is now the mass matrix arising from the discretisation of the operator $\mathbf{v} \longmapsto \mu\alpha^{-1}\mathbf{v}$, B^T is the discrete divergence operator, B its transpose and $\mathbf{G}$ the forcing term coming from the inhomogeneous Dirichlet condition applied to u. This system is in saddle point form and the system matrix therefore indefinite, i.e. it has negative and positive eigenvalues. In this thesis we will be motivated by the fact that in 2D system (1.3) can be decoupled into separate equations determining $\mathbf{V}$ and $\mathbf{U}$ and moreover that the equation for $\mathbf{V}$ is closely related to the discretisation of the problem

$$-\boldsymbol{\nabla}.\mu\alpha^{-1}\boldsymbol{\nabla}u = f \quad \text{on } \Omega \tag{1.4}$$

by standard finite elements (see Scheichl [116]).

We will return to the mixed finite element case in Chapter 7, but for most of the thesis will concentrate on a discretisation of

$$-\boldsymbol{\nabla}.\mu^{-1}\alpha\boldsymbol{\nabla}u = f \quad \text{on } \Omega \tag{1.5}$$

subject to suitable boundary conditions and using standard finite elements.

The discretisation of system (1.5) leads to a linear system of the form

$$A(\alpha)\mathbf{x} = \mathbf{y}, \tag{1.6}$$

where $A(\alpha)$ is a (sparse) symmetric positive definite matrix with entries that depend on the permeability α.

System (1.6) can be very hard to solve especially if the flow is through some strongly heterogeneous rock formations. These pose among others the following three problems:

(i) In heterogeneous media the permeability α can vary strongly over short distances. Changes of this coefficient by a factor of 10^9 or more on Ω are quite common in standard rock formations.

(ii) It is difficult to capture the heterogeneity of the medium. If one does so, the problem therefore has to be discretised on a very fine mesh, which means that the system matrix $A(\alpha)$ in (1.6) is typically of size ranging from $10^5 \times 10^5$ to $10^9 \times 10^9$.

(iii) The permeability can only be measured at a relatively small number of points. Away from these points however we have to deal with a significant uncertainty of α.

Problems (ii) and (iii) can be tackled by modeling the permeability field α as a stochastic spatial field, i.e. we identify one random variable with every point $\mathbf{x} \in \Omega$. The dependence of pairs of these random variables at points $\mathbf{x}, \mathbf{y} \in \Omega$ is described by a so called covariance function $\Sigma(\mathbf{x}, \mathbf{y})$. The strength of this dependence is expressed by a parameter λ, the correlation length of the field, where large λ (relative to the size of Ω) mean strongly correlated (smooth) fields, while small λ mean weakly correlated (rough) fields. We then solve system (1.6) for multiple realisations of the random field and compute statistical properties of u (or u and $\mathbf{v}$ when solving system (1.1,1.2)) using some Monte Carlo method.
Due to the large size of the system matrix, $A(\alpha)$, it is (especially for three dimensional domains Ω) difficult to solve system (1.6) directly even by exploiting the sparsity pattern of the matrix in sophisticated direct methods. Typically then Krylov subspace methods, like conjugate gradients (cg, see Appendix B), are used to solve it. It can be shown that for strongly heterogeneous fields (see (i)) $A(\alpha)$ is extremely ill-conditioned, i.e. its condition number $\kappa(A(\alpha))$ is large. In fact $\kappa(A(\alpha))$ can in the worst case depend linearly on the maximum ratio of α on Ω, which may be of order 10^9 or more for typical permeability fields of rock formations. As we will show theoretically and by numerical results this can cause the iterative method to converge very slowly for problems with strongly heterogeneous fields.

However if we multiply (1.6) by a matrix M^{-1}, then the solution of the resulting system

$$M^{-1}A(\alpha)\mathbf{x} = M^{-1}\mathbf{y}, \tag{1.7}$$

is still a solution of (1.6). The matrix M^{-1} is called a preconditioner of the system. However the convergence of the iterative method applied to the new system (1.7) now depends on $\kappa\left(M^{-1}A(\alpha)\right)$ instead of $\kappa\left(A(\alpha)\right)$. We therefore ideally want to choose M^{-1} in such a way that

- $\kappa\left(M^{-1}A(\alpha)\right) = \mathcal{O}(1)$,
- $\mathbf{z} := M^{-1}\mathbf{x}$ is cheap to compute for any vector $\mathbf{x}$.

The second condition is necessary, as this matrix-vector multiplication has to be performed in each iteration step of the Krylov method. Preconditioners which satisfy both these conditions are often very hard to find but essential, if we want to solve large systems in practice.

In this thesis we focus on preconditioners that are based on a decomposition of the domain Ω into nonoverlapping subdomains Ω_j, $j = 1, ..., p$. We then extend each subdomain Ω_j to an overlapping subdomain $\widetilde{\Omega_j}$ and solve a local problem on it. Furthermore we formulate and solve a problem on the coarse grid consisting of the subdomains Ω_j. To compute approximations of the values on the fine grid from the solution on the coarse grid in one subdomain Ω_j, the classical approach is to interpolate linearly between the values at the corner nodes of Ω_j. We will examine the dependence of this preconditioner on the permeability field α.

Next we will introduce another finite element space on the coarse grid, the so called space of multiscale finite elements (MsFEs). The functions in this space differ from the functions used for linear interpolation in that they depend on the underlying permeability field. Using the MsFE-basis functions in the coarse grid solves, a new domain decomposition preconditioner will be introduced and details of its performance will be given by theory and extensive numerical results for one-, two- and three-dimensional domains Ω for two media problems and for random permeability fields. We will study how its behaviour depends on the variance of the underlying field, the location of extreme values of the coefficients and the overlap between the subdomains $\widetilde{\Omega_j}$.

1.2 The main achievements of the thesis

The main achievements of this thesis include the following.

(i) Condition number estimates for the stiffness matrices arising from the groundwater flow problem are given. It is shown that the problem can be very ill-conditioned for heterogeneous media.

(ii) For the condition numbers of these stiffness matrices expectation number estimates in the general random field case (using simple norm inequalities) and estimates for the probability distribution of the condition number for fields of Ornstein-Uhlenbeck type (using Poisson clumping heuristics) are derived. It is shown that these are sharp in the 1D case and of high quality for higher dimensions.

(iii) Additive Schwarz preconditioners with linear coarsening for groundwater flow problems are studied. It is proved that a sharp upper bound of the condition number of the system preconditioned with them only depends on ratios of permeability values within the subdomains compared to global ratios in the unpreconditioned case. This can mean a clear improvement for fields with strong covariance, i.e. for fields in which points distanced far apart are correlated, and can therefore reduce the number of conjugate gradient iterations significantly.

(iv) However the problem can still be very ill-conditioned for fields with strong variance and/or short correlation length. Therefore a new coarsening technique based on multiscale finite elements is developed. Several boundary conditions for the construction of the coarse grid basis functions (used for the new preconditioners) are discussed and compared.

(v) By considering some simple model problems we develop bounds for the condition numbers of preconditioned stiffness matrices when applying additive Schwarz preconditioners with MsFE-coarsening. Some ideas for general fields are given. The significance of how the location of the extreme values of the coefficients affects the convergence of the method is researched.

(vi) It is shown that the new preconditioner is the exact inverse of the stiffness matrix in the 1D case. The same is also proved for an extended version of it in higher dimensional cases and non-overlapping subdomains.

(vii) Additive Schwarz preconditioners without coarse grid, with linear and with multiscale coarsening are compared. The theory is then supported by extensive

numerical experiments in 1D, 2D and 3D for random fields of general and of Ornstein-Uhlenbeck type, as well as for two media problems.

(viii) The newly developed preconditioner is applied to a mixed finite element formulation of the groundwater flow problem. Numerical results in 2D are given.

1.3 The structure of the thesis

This thesis consists of six main chapters which try to focus on different aspects of the analysis of the iterative solution of groundwater flow problems. Each chapter contains a preamble with a motivation of the subsequent ideas and concludes with a summary in which we try to set the new results and observations into perspective with the rest of the thesis. This structure will hopefully help the reader to see the overall picture of the work and should also allow the use of this thesis for later reference.

Chapter 2 introduces the problem considered, a second-order partial differential equation with strongly varying coefficients. Its meaning in groundwater flow applications as well as a weak formulation are given. We then derive and describe for 2D problems a mixed formulation of the problem using first order Raviart-Thomas finite elements. Using a divergence free reduction we can separate velocity and pressure, which will enable us later to use the preconditioners, which are developed in Chapters 5 to 7, for this problem.

To model the underlying permeability fields for groundwater flow it is quite common to use random fields of Ornstein-Uhlenbeck type. Their properties are discussed in **Chapter 3**. We then find upper bounds for the expectation of the maxima of general random fields based on simple norm inequalities as well as approximations for the probability distributions of maxima for random fields of Ornstein-Uhlenbeck type based on Poisson clumping heuristics. These theoretical bounds are used in Chapters 4 and 5 to get estimates for the unpreconditioned and preconditioned stiffness matrices for Poisson's equation with underlying random permeability fields.

Chapter 4 shows why the problems considered are ill-conditioned for strongly varying permeability fields. We give estimates in the deterministic as well as in the random media case based on the results from Chapter 3.

In **Chapter 5** we introduce additive Schwarz preconditioners without coarse grid as well as with linear coarsening. We find new bounds which show the dependence of the condition number of the stiffness matrix preconditioned using this domain decomposition method. Furthermore we give numerical results which show the strong increase of these condition numbers with increasing variance of the underlying permeability field.

Chapter 6 deals with exactly this problem by introducing a new coarsening operator

based on multiscale finite elements. The importance of good boundary conditions for the construction of the multiscale basis functions is discussed and supported by a long series of numerical results. By considering some simple model problems we give a full convergence analysis for certain special cases and ideas of an analysis for general fields and general overlap. We show why for one-dimensional problems this preconditioner and for higher dimensional problems an extension of it delivers an exact inverse of the stiffness matrix.

Having discussed the theoretical properties of the additive Schwarz preconditioners without coarse grid, as well as with linear and multiscale coarsening, we then compare their quality in terms of the condition numbers of the preconditioned stiffness matrices as well as their effects on computation times, when used in iterative solvers in **Chapter 7**. Numerical experiments are performed for a two media case as well as for random fields in 1D, 2D and 3D. The preconditioners are finally applied to mixed formulations of the groundwater flow problem.

In **Appendix A** some basic results from Linear Algebra on eigenvalues and eigenvectors for preconditioned and unpreconditioned systems are given. **Appendix B** explains the most important aspects of the iterative method used for our numerical tests, the conjugate gradient method, and shows the importance of small condition numbers of the system matrix for a fast convergence of the method. Throughout the thesis we try to stick with a fairly standard notation, which is summarised in **Appendix C** to help the reader to find his/her way through this work.

Chapter 2

Problem formulation

"If all the world's water fit into a bathtub, the portion of it that could be used sustainably in any given year would barely fill a teaspoon."

— World Resources Institute, 1993 —

"In questions of science, the authority of a thousand is not worth the humble reasoning of a single individual."

— Galileo Galilei (Italian natural Philosopher, Astronomer and Mathematician who made fundamental contributions to the development of the scientific method and to the sciences of motion, astronomy and strength of materials. 1564-1642) —

2.1 Derivation of Groundwater Flow equations

Groundwater is and will become even more in the next few decades a scarce and highly important resource. It has been extracted intensively in some regions (India, California, etc.) and this has lead to a lowering of the water table, beyond the reach of existing wells. This means that better technology, but also better mathematical models to compute the movement of groundwater have become very essential to guarantee sufficient supply of fresh water for an increasing population, especially in developing regions. Conflicts between countries about water might otherwise soon become reality. Groundwater is flowing in so called *aquifers*, underground layers of water-bearing rock

or permeable mixtures of unconsolidated materials (sand, gravel, etc.). It flows between porous spaces and can naturally be recharged through the surface. It is estimated that the amount of water underground exceeds that of fresh water on the surface by fifty times.
Flow models for groundwater can also be used for pollution modeling, for example to simulate the transport of radionuclides in flowing groundwater. Furthermore they help to understand the underground flow of oil and other natural resources, to be able to recover as much of them as possible at the lowest possible cost. This is important as for example drilling one hole for the recovery of oil offshore costs many million US Dollars.

Before giving any technical details, this section will serve to set the scene and will derive the underlying PDE model for the groundwater flow problem considered.
A literature review describing the use of groundwater flow models can be found in Prickett [109]. The fundamental groundwater physics are described by Bennett [10], who covers definitions and general concepts (Part I), Darcy's law (Part II) and applications of Darcy's law (Part III). Mercer and Faust give a good idea of how groundwater flow can be described with the help of mathematical models in their three paper series giving a general overview (Mercer and Faust [100]), introducing mathematical models (Mercer and Faust [101]) and possible applications (Mercer and Faust [102]). *Fluid flow models* are mainly based on two fundamental principles, the *conservation of mass* and the *conservation of momentum.*
The first classical equation governing this application is *Darcy's law* (see Darcy [34]). This law was first formulated by *Henry Darcy* after obtaining data from experiments on the flow of water through sand beds. Further details on Darcy's law can also be found in Bennett [10], Freeze and Cherry [57], Hubbert [77] and Wang and Anderson [131]. It expresses the conservation of momentum and is analogous to Fourier's law in the field of heat conduction, Ohm's law in the field of electrical networks, or Fick's law in diffusion theory.
For this equation let α denote the *intrinsic permeability* (units of m^2 or Darcy, where 1 Darcy $\approx 10^{-12}m^2$). The intrinsic permeability is generally a tensor, but we will only consider the scalar valued case in this thesis. Furthermore let v be the *specific discharge or Darcy velocity*, u_R the *residual pressure* and μ the *dynamic viscosity of the fluid.*

Darcy's law is then given by

$$\mathbf{v} = -(\alpha/\mu)\,\nabla u_R. \tag{2.1}$$

In the following we will assume for simplicity that $\mu = 1$. The *actual pressure* is $u_R - \rho g z$, where z is the *fluid height*, ρ is the density and g is the *acceleration due to*

gravity.

From this equation, we can see among others:

- If there is no change in (hydraulic) pressure, no flow occurs.
- If change in pressure occurs, the flow is from areas with high pressure to areas with low pressure.
- The greater the pressure gradient, the stronger the flow.

In practice one can show that Darcy's law delivers good results for slow, viscous flows, i.e. for flow regimes with *Reynold's number* $Re < 1$.

Next to Darcy's law for steady-state conditions, continuity requires that the amount of water that flows into a representative elemental volume equals the amount flowing out of it. Let us therefore also assume the following:

- Water is incompressible. (Otherwise it could be compressed and stored in the elemental volume. Then mass would still be conserved, but volume would not.)
- The elemental volume contains no sources or sinks, i.e. no water is allowed to be added or removed within the elemental volume.

Let us now consider an elemental cube with sides of length Δx, Δy and Δz. Its volume is therefore given by $V := \Delta x \Delta y \Delta z$. For the mass balance we now sum the results from each component direction. Therefore v_y for example represents the volume rate of flow per unit area through the left face of the cube. The discharge through the left and right faces is then given by the product of the flow rate per unit area times the area $\Delta x \Delta z$ of these faces. At the right face the discharge differs by $(\partial v_y/\partial y)\,\Delta y$. The net discharge rate in the y-direction is then given by $(\partial v_y/\partial y)\,\Delta y\,(\Delta x \Delta z) = (\partial v_y/\partial y)\,\Delta V$. Similarly the net change in the discharge rate in the x-direction is $(\partial v_x/\partial x)\,\Delta V$, the net change in the discharge rate in the z-direction is $(\partial v_z/\partial z)\,\Delta V$. For mass conservation the sum of these three net changes now has to equal 0. Dividing by ΔV we obtain the *continuity equation for steady state conditions*

$$\operatorname{div}(\mathbf{v}) = \frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} + \frac{\partial v_z}{\partial z} = 0. \tag{2.2}$$

Assuming $\mu = 1$ and combining steady state equation (2.2) with Darcy's law (2.1)

finally leads to the *Laplace equation*

$$\operatorname{div}(\alpha(\mathbf{x})\nabla u_R(\mathbf{x})) = 0 \text{ for } \mathbf{x} \in \Omega, \tag{2.3}$$

with suitable boundary conditions, that will be discussed in the next paragraph, and under the following assumptions:

(i) The aquifer material is incompressible.

(ii) The water is of constant density.

(iii) The external pressure on the aquifer is constant.

(iv) The groundwater is flowing slowly (Reynold's number less than 1).

(v) The hydraulic permeability α is an isotropic scalar.

For ease of notation we now replace u_R by u in equations (2.1) and (2.3). This equation is a *second order elliptic partial differential equation.*
Boundary conditions for the problem can for example be of the form

$$\xi u + \zeta\,(\alpha\nabla u)\,.\boldsymbol{\nu} = g \text{ on } \Gamma, \tag{2.4}$$

where $\boldsymbol{\nu}$ denotes the outwards pointing normal to the boundary Γ and $\xi, \zeta \in \mathbb{R}$. Combining (2.3) and (2.4) gives a *boundary value problem.* If ζ is zero, the boundary condition is said to be of *Dirichlet type*; if ξ is zero, the boundary condition is said to be of *Neumann type.* A third possibility is $\xi \neq 0$ and $\zeta = 0$ on one part and $\zeta \neq 0$ and $\xi = 0$ on the other part of the boundary. In this case we talk of *mixed* boundary conditions. If both $\xi \neq 0$ and $\zeta \neq 0$ the problem has so called *Robin boundary conditions.*

As a simple model problem, we can for example consider the unit square $\Omega = (0,1)^2$ in two and the unit cube $\Omega = (0,1)^3$ in three dimensions. Possible boundary conditions for a 2D model problem are given in Figure 2-1.
More generally we assume that the boundary Γ of Ω can be split into two parts Γ_D (on which Dirichlet conditions hold, i.e. $\xi \neq 0$ and $\zeta = 0$) and Γ_N (on which Neumann conditions hold, i.e. $\zeta \neq 0$ and $\xi = 0$), such that $\Gamma = \Gamma_D \cup \Gamma_N$ and $\Gamma_D \cap \Gamma_N = \emptyset$.
Replacing the zero right hand side of (2.3) by a more general function f, we next consider the following problem.

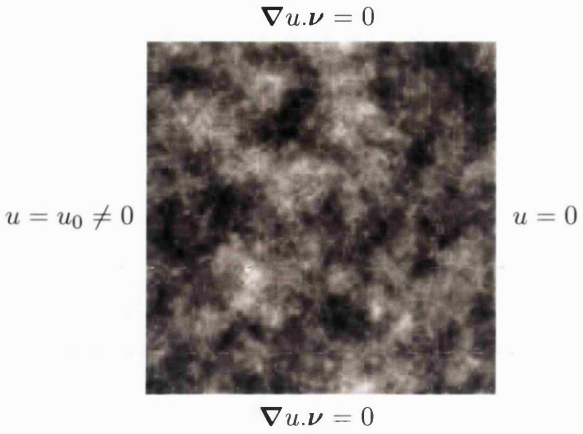


Figure 2-1: Boundary conditions for the groundwater flow problem on $(0,1) \times (0,1)$ (Showing in grey scale possible variations of the permeability α, which we will discuss in more detail in Chapter 3)

$$-\text{div}\,(\alpha \nabla u(\mathbf{x})) = f(\mathbf{x}) \text{ for } \mathbf{x} \in \Omega, \tag{2.8}$$

with

$$u(\mathbf{x}) = g(\mathbf{x}) \text{ for } \mathbf{x} \in \Gamma_D \tag{2.9}$$

and

$$\alpha \nabla u(\mathbf{x}).\nu = 0 \text{ for } \mathbf{x} \in \Gamma_N. \tag{2.10}$$

By discussing this problem, we treat a whole class of groundwater problems, as they are special cases of it.
In the following we will consider the problem (2.8, 2.9, 2.10), but some of the theory will be presented in the case of the simpler *Dirichlet problem*:

$$-\text{div}\,(\alpha \nabla u(\mathbf{x})) = f(\mathbf{x}) \text{ for } \mathbf{x} \in \Omega, \tag{2.13}$$

with

$$u(\mathbf{x}) = 0 \text{ for } \mathbf{x} \in \Gamma, \tag{2.14}$$

on $\Omega = (0,1)^d$, where d is the dimension of the problem.

2.2 H^1 finite element approximation

2.2.1 Weak formulation

If the solution u of the boundary value problem (2.8, 2.9, 2.10) is sufficiently smooth, it is known as a *classical solution.* For a problem of this form on a convex domain Ω with pure Dirichlet boundary conditions, u is a classical solution only if it has a second derivative that is continuous on the whole of the open domain Ω, i.e. if $u \in C^2(\Omega) \cap C^0(\bar{\Omega})$, where $\bar{\Omega} := \Omega \cup \Gamma$. If however for example α and/or f are/is not continuous on the whole of Ω, there can not exist a classical solution to the problem. This situation can arise for perfectly reasonable mathematical problems and it is often important to find an alternative formulation of the problem that is less restrictive in terms of the admissible data, the so called *weak formulation.*
Let us therefore define the spaces

$$L_2(\Omega) := \left\{ u : \Omega \to \mathbb{R} \mid \int_\Omega u^2 < \infty \right\},$$

$$H^1(\Omega) := \left\{ u : \Omega \to \mathbb{R} \mid u, \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y} \in L_2(\Omega) \right\},$$

$$\mathcal{V} := \left\{ v : \Omega \to \mathbb{R} \mid v \in H^1(\Omega), v = 0 \text{ on } \Gamma_D \right\}.$$

as well as

$$\mathcal{V}_g := \left\{ v : \Omega \to \mathbb{R} \mid v \in H^1(\Omega), v = g \text{ on } \Gamma_D \right\}.$$

Next define the bilinear form

$$(u, v)_{H^1(\Omega),\alpha} := \int_\Omega \alpha\,(\mathbf{x})\, \boldsymbol{\nabla} u(\mathbf{x}).\boldsymbol{\nabla} v(\mathbf{x}) \mathrm{d}\mathbf{x},$$

the semi-norm

$$|u|_{H^1(\Omega),\alpha} = \int_\Omega \alpha\,(\mathbf{x})\, \boldsymbol{\nabla} u(\mathbf{x}).\boldsymbol{\nabla} u(\mathbf{x}) \mathrm{d}\mathbf{x}$$

and

$$L(v) := \int_\Omega f(\mathbf{x}) v(\mathbf{x}) \mathrm{d}\mathbf{x} + \int_{\Gamma_N} g(\mathbf{x}) v(\mathbf{x}) \mathrm{d}\mathbf{x}.$$

Then the *weak formulation* of (2.8,2.9,2.10) is given by:

Find $u \in \mathcal{V}_g$ such that $(u,v)_{H^1(\Omega),\alpha} = L(v), \quad \forall v \in \mathcal{V}.$ (2.16)

2.2.2 Galerkin finite element approximation

Next we would like to find an approximation u_h of u for the boundary value problem (2.8, 2.9, 2.10), resp. its weak formulation (2.16) in a finite dimensional subspace of $\mathcal{V}$. Let therefore $\bar{\Omega}$ denote the closure of domain Ω.

DEFINITION 2.1

(i) A simplical triangulation $\mathcal{T}(\bar{\Omega})$ is a partitioning of $\bar{\Omega} \in \mathbb{R}^n$ into closed intervals (in 1D), triangles (in 2D), resp. tetrahedra (in 3D), such that

- $\bar{\Omega} = \bigcup_{\tau \in \mathcal{T}} \tau$,
- $\tau \cap \tau' = \begin{cases} \text{either empty} \\ \text{or a common } \textbf{vertex} \text{ of } \tau \text{ and } \tau' \\ \text{or (for } n \geq 2) \text{ a common } \textbf{edge} \text{ of } \tau \text{ and } \tau' \\ \text{or (for } n = 3) \text{ a common } \textbf{face} \text{ of } \tau \text{ and } \tau' \end{cases}$

for all $\tau \neq \tau' \in \mathcal{T}(\bar{\Omega})$.

The $\tau \in \mathcal{T}$ are called elements. Let h_τ be the diameter of $\tau \in \mathcal{T}$, let $h := \max_{\tau \in \mathcal{T}} h(\tau)$ and let ρ_τ be the diameter of the largest circle (or ball for $n = 2$) contained within τ. Furthermore denote $\mathcal{T}^h(\bar{\Omega})$ instead of $\mathcal{T}(\bar{\Omega})$ if we study the assymptotics for $h \to 0$.

*(ii) A family of triangulations $\{\mathcal{T}^h(\bar{\Omega})\}$ is called **(shape) regular**, if there exists a constant C (independent of h) such that*

$$\frac{h_\tau}{\rho_\tau} \leq C. \tag{2.17}$$

*(iii) A family of finite element subdivisions $\{\mathcal{T}^h(\bar{\Omega})\}$ is called **quasi-uniform** if it is regular and if there is a constant C_1 (independent of h), such that for all $\tau \in \mathcal{T}^h(\bar{\Omega})$ we have $h_\tau \geq C_1 h$.*

(iv) A ***finite element*** *in* $\mathbb{R}^n$ *is a triple* $(\tau, \mathcal{P}, \mathcal{L})$ *where*

- τ *is a closed subset of* $\mathbb{R}^n$ *with a non empty interior and a Lipschitz-continuous boundary.*
- $\mathcal{P}$ *is a space of real-valued functions defined over the set* τ.
- $\mathcal{L}$ *is a finite set of linearly independent linear forms* p_i, $1 \leq i \leq n$, *defined over the space* $\mathcal{P}$ *(in order to avoid ambiguities, the form* p_i *need to be defined over a larger space). By definition, it is assumed that the set* $\mathcal{L}$ *is unisolvent in the following sense: given any real scalars* s_i, $1 \leq i \leq n$, *there exists a unique function* $\phi \in \mathcal{P}$ *which satisfies*

$$p_i(\phi) = s_i, \quad 1 \leq i \leq n. \tag{2.18}$$

Consequently there exist functions $\phi_i \in P$, $1 \leq i \leq n$, *which satisfy*

$$\phi_i(p_i) = \delta_{ij}, \quad 1 \leq j \leq n. \tag{2.19}$$

Therefore we have

$$\forall \phi \in \mathcal{P}, \quad \phi = \sum_{i=1}^{n} p_i(\phi)\, \phi_i. \tag{2.20}$$

(v) The linear forms p_i, $1 \leq i \leq n$, *are called the* ***degrees of freedom*** *of the finite element. The functions* ϕ_i, $1 \leq i \leq n$ *are called the* ***basis functions*** *of the finite element.*

Assume now that we have a family of meshes satisfying (i) to (v).

Let us then introduce the finite dimensional subspaces

$$\mathcal{V}^h := \left\{ v_h \in C(\Omega) \quad | \quad v_h|_\tau \text{ linear } \forall \tau \in \mathcal{T}^h(\bar{\Omega}) \text{ and } v = 0 \text{ on } \Gamma_D \right\}$$

and

$$\mathcal{V}_g^h := \left\{ v_h \in C(\Omega) \quad | \quad v_h|_\tau \text{ linear } \forall \tau \in \mathcal{T}^h(\bar{\Omega}) \text{ and } v = g \text{ on } \Gamma_D \right\}.$$

If $\mathcal{N}^h(\bar{\Omega})$ is the set of nodes $\mathbf{x}_i^h$ of the triangulation $\mathcal{T}^h(\bar{\Omega})$ (including all nodes on the boundary Γ), then a basis of $\mathcal{V}^h$ is given by the so called *hat functions* $\{\phi_i : \mathbf{x}_i^h \in \mathcal{N}^h(\bar{\Omega})\}$, where ϕ_i is defined to be piecewise linear with respect to this mesh and for which

$\phi_i(\mathbf{x}_j^h) := \delta_{ij}$, where δ_{ij} denotes the Kronecker delta.
Hence the *approximate weak form* of (2.16) is given by:

Find $u_h \in \mathcal{V}_g^h$ such that $(u_h, v_h)_{H^1(\Omega),\alpha} = L(v_h), \quad \forall v_h \in \mathcal{V}^h.$ (2.22)

We can then write

$$u_h = \sum_{\mathbf{x}_i^h \in \mathcal{N}^h(\bar{\Omega})} U_i \phi_i, \tag{2.23}$$

with coefficients $U_i \in \mathbb{R}$, where values of U_i are known at the nodes on Γ_D.

Substituting this into the approximate weak form (2.22), we get, since $\phi_i \in \mathcal{V}^h$ for all i with $\mathbf{x}_i \in \mathcal{N}^h(\Omega)$,

$$\sum_{\mathbf{x}_i^h \in \mathcal{N}^h(\bar{\Omega})} (\phi_i, \phi_j)_{H^1(\Omega),\alpha} U_i = L(\phi_j), \quad \mathbf{x}_j^h \in \mathcal{N}^h(\bar{\Omega}) \backslash \Gamma_D \tag{2.24}$$

or equivalently

$$\sum_{\mathbf{x}_i^h \in \mathcal{N}^h(\bar{\Omega})} (\phi_i, \phi_j)_{H^1(\Omega),\alpha} U_i = L(\phi_j) - \sum_{\mathbf{x}_i^h \in \mathcal{N}^h(\bar{\Omega}) \cap \Gamma_D} (\phi_i, \phi_j)_{H^1(\Omega),\alpha} g\left(\mathbf{x}_i^h\right), \tag{2.25}$$

for all j such that $\mathbf{x}_j^h \in \mathcal{N}^h(\bar{\Omega}) \backslash \Gamma_D$.

Setting $A_{ij} := (\phi_i, \phi_j)_{H^1(\Omega),\alpha}$ and $f_i := L(\phi_i)$, this leads to the linear system

$$A\mathbf{U} = \mathbf{f}, \tag{2.26}$$

where $A = A(\alpha)$ is the so called *stiffness matrix* and $\mathbf{f}$ is the so called *load vector* of the discretised system. Here $A_{ij} := (\phi_i, \phi_j)_{H^1(\Omega),\alpha}$ is only nonzero, if $\mathbf{x}_i^h$ and $\mathbf{x}_j^h$ are nodes of the same element.
Because of the strong fine scale variations of the underlying medium, one often has to choose the maximum element diameter h very small with respect to the size of the domain to resolve these structures sufficiently. This means that, especially for the three dimensional problems, systems (2.26) with several million unknowns are not uncommon. This however means that it is very expensive and in many cases impossible

to solve the problem directly. *Iterative methods* are therefore essential and will be discussed in detail in the following chapters.

2.3 Mixed finite element approximation

2.3.1 Mixed formulation

In many applications the velocity $\mathbf{v} = \alpha \nabla u$, where we assume again that the permeability $\mu = 1$ and where α denotes the permeability, is of greater interest than the pressure u. One therefore often has to investigate the saddle point problem arising from (2.8) to (2.10), in the unknowns u (pressure) and $\mathbf{v}$ (velocity).
With $\Omega \subset \mathbb{R}^d$, one can now introduce the function space

$$H(\mathrm{div}, \Omega) := \left\{ \mathbf{w} \in (L_2(\Omega))^d : \mathrm{div}(\mathbf{w}) \in L_2(\Omega) \right\} \tag{2.27}$$

and the inner product

$$(\mathbf{v}, \mathbf{w})_{H(\mathrm{div},\Omega)} := \int_\Omega (\mathbf{v}.\mathbf{w} + \mathrm{div}(\mathbf{v}) \cdot \mathrm{div}(\mathbf{w}))\, \mathrm{d}\mathbf{x}. \tag{2.28}$$

The space $L_2(\Omega)$ is equipped with its usual inner product

$$(u, s)_{L_2(\Omega)} = \int_\Omega u(\mathbf{x}) s(\mathbf{x}) \mathrm{d}\mathbf{x} \tag{2.29}$$

and with the induced norm $\|\cdot\|_{L_2(\Omega)}$.
We can then formulate (2.8, 2.9, 2.10) as a saddle point problem in the two distinct spaces

$$\mathcal{R} := H_{0,N}(\mathrm{div}, \Omega) = \{ \mathbf{v} \in H(\mathrm{div}, \Omega) \mid \mathbf{v}.\boldsymbol{\nu} = 0 \text{ on } \Gamma_N \}$$

and $\mathcal{S} := L_2(\Omega)$.
Let us therefore consider the (groundwater flow) system (as in (1.1)):

$$\alpha^{-1}\mathbf{v} + \nabla u = 0, \tag{2.30}$$

$$\mathrm{div}(\mathbf{v}) = 0, \tag{2.31}$$

subject to (2.9,2.10).
Multiplying equation (2.30) by an arbitrary function $\mathbf{r} \in \mathcal{R}$ and equation (2.31) by an

arbitrary function $s \in \mathcal{S}$ and using

$$(\nabla u, \mathbf{r})_{L_2(\Omega)} = -(u, \mathrm{div}(\mathbf{r}))_{L_2(\Omega)} + \langle g, \mathbf{r}.\boldsymbol{\nu} \rangle_{L_2(\Gamma_D)},$$

where $\langle g, \mathbf{r}.\boldsymbol{\nu} \rangle_{L_2(\Gamma_D)} := \int_{\Gamma_D} g\mathbf{r}.\boldsymbol{\nu} \mathrm{d}\mathbf{x}$, we obtain the weak form of (2.30, 2.31), given by

Find a solution $(\mathbf{v}, u) \in \mathcal{R} \times \mathcal{S}$ satisfying

$$\begin{aligned} \left(\alpha^{-1}\mathbf{v}, \mathbf{r}\right)_{L_2(\Omega)} - \left(\mathrm{div}\left(\mathbf{r}\right), u\right)_{L_2(\Omega)} &= \langle g, \mathbf{r}.\boldsymbol{\nu} \rangle_{L_2(\Gamma_D)}, \quad \forall \mathbf{r} \in \mathcal{R}, & (2.34) \\ \left(\mathrm{div}\left(\mathbf{v}\right), s\right)_{L_2(\Omega)} &= 0, \quad \forall s \in \mathcal{S}. & (2.35) \end{aligned}$$

It can be shown that this problem has a unique solution (see Raviart and Thomas [111] and Brezzi and Fortin [19]).

To find a finite element discretisation for this mixed formulation, we now choose finite-dimensional subspaces $\mathcal{R}_h \subset \mathcal{R}$ and $\mathcal{S}_h \subset \mathcal{S}$ and then try to solve the following problem.

Find a solution $(\mathbf{v}_h, u_h) \in \mathcal{R}_h \times \mathcal{S}_h$ satisfying

$$\begin{aligned} \left(\alpha^{-1}\mathbf{v}_h, \mathbf{r}_h\right)_{L_2(\Omega)} - \left(\mathrm{div}\left(\mathbf{r}_h\right), u_h\right)_{L_2(\Omega)} &= \langle g, \mathbf{r}_h.\boldsymbol{\nu} \rangle_{L_2(\Gamma_D)}, \quad \forall \mathbf{r}_h \in \mathcal{R}_h, & (2.38) \\ \left(\mathrm{div}\left(\mathbf{v}_h\right), s_h\right)_{L_2(\Omega)} &= 0, \quad \forall s_h \in \mathcal{S}_h. & (2.39) \end{aligned}$$

Let us now introduce the bilinear forms

$$w(\mathbf{v}, \mathbf{r}) := \left(\alpha^{-1}\mathbf{v}, \mathbf{r}\right)_{L_2(\Omega)} \quad \text{and} \quad b(\mathbf{r}, s) := -\left(\mathrm{div}\left(\mathbf{r}\right), s\right)_{L_2(\Omega)}$$

as well as the linear functional

$$G(\mathbf{r}) := \langle g, \mathbf{r}.\boldsymbol{\nu} \rangle_{L_2(\Gamma_D)}.$$

Then (2.38, 2.39) can be written:

Find a solution $(\mathbf{v}_h, u_h) \in \mathcal{R}_h \times \mathcal{S}_h$ satisfying

$$w(\mathbf{v}_h, \mathbf{r}_h) + b(\mathbf{r}_h, u_h) = G(\mathbf{r}_h), \quad \forall \mathbf{r}_h \in \mathcal{R}_h, \tag{2.40}$$

$$b(\mathbf{v}_h, s_h) = 0, \quad \forall s_h \in \mathcal{S}_h. \tag{2.41}$$

It was shown by Brezzi and Fortin [19] that this problem has a unique solution if $\operatorname{div}(\mathcal{R}_h) = \mathcal{S}_h$.

If we now choose bases $\{\mathbf{r}_1, ..., \mathbf{r}_{n_r}\}$ for $\mathcal{R}_h$ and $\{s_1, ...s_{n_s}\}$ for $\mathcal{S}_h$, where n_r denotes the dimension of $\mathcal{R}_h$ and where n_s denotes the dimension of $\mathcal{S}_h$ and expressing

$$\mathbf{v}_h := \sum_{i=1}^{n_r} V_i \mathbf{r}_i \quad \text{and} \quad u_h := \sum_{i=1}^{n_s} U_i s_i$$

in terms of these bases, the weak formulation (2.40, 2.41), then reduces to the indefinite linear equation system

$$\begin{bmatrix} W(\alpha) & B \\ B^T & 0 \end{bmatrix} \begin{pmatrix} \mathbf{V} \\ \mathbf{U} \end{pmatrix} = \begin{pmatrix} \mathbf{G} \\ \mathbf{0} \end{pmatrix}. \tag{2.42}$$

Here the *mass matrix* $W(\alpha)$ is given by $W_{i,j}(\alpha) := w(\mathbf{r}_i, \mathbf{r}_j)$, the *discrete gradient* B by $B_{i,j} := b(\mathbf{r}_i, s_j)$ and the vector $\mathbf{G}$ by $G_i := G(\mathbf{r}_i)$.

2.3.2 Raviart-Thomas discretisation

Let us now restrict to two dimensional domains Ω and to the lowest order *Raviart-Thomas space* $\mathcal{R}_h$. This finite element space was introduced by Raviart and Thomas in two dimensions (see Raviart and Thomas [111]) and then extended to three dimensions by Nédélec (see Nédélec [104]).

We define the space $\mathcal{R}_h$ to consist of all vector-valued functions $\mathbf{r}_h \in \mathcal{R}$, such that for all $\tau \in \mathcal{T}^h(\bar{\Omega})$, there exist scalars c_τ, d_τ and e_τ, such that

$$\mathbf{r}_h(\mathbf{x}) = \begin{pmatrix} c_\tau \\ d_\tau \end{pmatrix} + e_\tau \mathbf{x}, \quad \forall \mathbf{x} \in \tau. \tag{2.43}$$

After this choice for $\mathcal{R}_h$, we also have to choose $\mathcal{S}_h$. To guarantee existence and uniqueness of the solution of system (2.42) (and with $\operatorname{div}(\mathcal{R}_h) = \mathcal{S}_h$) we can therefore choose $\mathcal{S}_h$ as the space of piecewise constant functions on Ω, with the basis chosen to

be the characteristic functions of each $\tau \in \mathcal{T}^h(\bar{\Omega})$.

2.4 Divergence free reduction

Let us now come back to system (2.42) and let $\mathbf{U}$ be the vector of pressures and $\mathbf{V}$ the vector of velocities at the freedoms in $\mathcal{R}_h$ and $\mathcal{S}_h$. We would like to reduce the indefinite coupled system (2.42) to a smaller symmetric positive definite system to be able to apply simple Krylov space methods as well as the preconditioners that we will develop in Chapters 4 to 6. The following method has been discussed by Scheichl [116, 117]. Earlier work on divergence-free elements was done by Thomasset [125, 126], Hecht [70, 71], Griffiths [61], Nédélec [105], Mack [94] and Ye and Anderson [133].

Let again n_r be the dimension of $\mathcal{R}_h$ and n_s be the dimension of $\mathcal{S}_h$. As shown by Brezzi and Fortin [19], the saddle point system (2.42) has got a unique solution $(\mathbf{V}, \mathbf{U}) \in \mathbb{R}^{n_r} \times \mathbb{R}^{n_s}$ for all $\mathbf{G} \in \mathbb{R}^{n_r}$. Furthermore since the right hand side of (2.42) is $\left(\mathbf{G}^T, \mathbf{0}^T\right)^T$, $\mathbf{V} \in \ker\left(B^T\right)$, the kernel of B^T.

The *decoupling procedure* consists of the following four steps:

Step 1: Find a basis $\left\{\mathbf{z}_1, ..., \mathbf{z}_{\mathring{n}}\right\}$ of $\ker\left(B^T\right)$ (as B^T has full rank, $\mathring{n} = n_r - n_s$).

Step 2: With this basis, the solution $\mathbf{V}$ can be written

$$\mathbf{V} = \sum_{j=1}^{\mathring{n}} \mathring{V}_j \mathbf{z}_j = Z^T \mathring{\mathbf{V}}, \tag{2.44}$$

for some $\mathring{\mathbf{V}} \in \mathbb{R}^{\mathring{n}}$, with Z being the $\mathring{n} \times n_r$ matrix with rows $\mathbf{z}_1^T, ...\mathbf{z}_{\mathring{n}}^T$. Furthermore since $ZB = \left(B^T Z^T\right)^T = 0$, we get after multiplying the first row of the system (2.42) by Z, that $\mathring{\mathbf{V}}$ is a solution of the spd system

$$\mathring{A}\mathring{\mathbf{V}} = \mathring{\mathbf{G}}, \tag{2.45}$$

with

$$\mathring{A} := ZWZ^T \text{ and } \mathring{\mathbf{G}} := Z\mathbf{G}. \tag{2.46}$$

Solving this decoupled spd system (2.45) is therefore equivalent to solving the coupled indefinite system (2.42). If we are only interested in the velocity $\mathbf{V}$, then our compu-

tations are already finished here.

Step 3: If we are also interested in the pressure $\mathbf{U}$, we also compute a complementary basis $\{\mathbf{z}_{\mathring{n}}, ..., \mathbf{z}_{n_r}\}$, such that span $(\mathbf{z}_1, ..., \mathbf{z}_{\mathring{n}}, ..., \mathbf{z}_{n_r}) = \mathbb{R}^{n_r}$.

Step 4: Given this basis, let Z' be the matrix with rows $\mathbf{z}_{\mathring{n}}, ..., \mathbf{z}_{n_r}$. Multiplying the first row of blocks of system (2.42) by Z', we see that $\mathbf{U}$ is the solution of the system

$$\left(Z'B\right)\mathbf{U} = Z'\left(\mathbf{G} - W\mathbf{V}\right). \tag{2.47}$$

To find a basis $\{\mathbf{z}_1, ..., \mathbf{z}_{\mathring{n}}\}$ of ker $\left(B^T\right)$, note that by the definition of B, this is equivalent to finding a basis $\mathring{\mathbf{v}}_1, ..., \mathring{\mathbf{v}}_{\mathring{n}}$ of the finite element space

$$\mathring{\mathcal{V}} := \{\mathbf{r}_h \in \mathcal{R}_h : \; b\left(\mathbf{r}_h, s_h\right) = 0 \text{ for all } s_h \in \mathcal{S}_h\}.$$

In 2D a construction of this basis in the lowest order case was given by Cliffe et al. [31] and Scheichl [116]. In the latter reference, also higher order Raviart-Thomas elements in 2D and the lowest order case in 3D were studied.

We will discuss the construction of the basis of ker $\left(B^T\right)$ in some more detail for a simple model problem on a unit square domain $\Omega = (0,1)^2$ with two Dirichlet boundaries and two Neumann boundaries in Section 7.5. Furthermore we will then give some numerical results for various permeability fields.

To conclude this chapter we merely mention that for this reduction the matrix $\mathring{A}(\alpha)$ in (2.45) has the structure

$$\mathring{A}(\alpha) = \begin{bmatrix} A(\alpha) & C(\alpha) \\ C^T(\alpha) & D(\alpha) \end{bmatrix}, \tag{2.48}$$

where $A(\alpha)$ is a certain symmetric positive definite minor of the stiffness matrix corresponding to the approximation of the bilinear form $(\phi, \psi) \longmapsto \left(\alpha^{-1}\nabla\phi, \nabla\psi\right)$ using continuous piecewise linear elements and $\left[C^T(\alpha) \quad D(\alpha)\right]$ is a low-dimensional border. (For a proof of this statement see Scheichl [116], Proposition 4.2.) Using bordered matrix techniques we only have to solve systems with the coefficient matrix $A(\alpha)$ (see also Section 7.5).

Furthermore, it can be shown that system (2.45) is in most cases significantly smaller than the original mixed system (2.42). Probably the main advantage of the reduction

however is, that $A(\alpha)$ in (2.48) is of the same form as the system matrix in (2.26). We can therefore use the same sort of preconditioners to solve this system iteratively, e.g. the domain decomposition preconditioners that will be discussed in Chapters 5 to 7. A 2D example for this mixed formulation as well as several numerical results for the additive Schwarz preconditioned groundwater flow problem are given in Section 7.5.
A more general discussion of Raviart-Thomas elements can be found in Brezzi and Fortin [19].

2.5 Summary

The main focus of this short chapter was on describing the class of groundwater flow problems that we consider in this thesis. We first derived the second order partial differential equation (2.3) based on Darcy flow and the conservation of mass. We then gave the weak form of problem (2.16) and its Galerkin approximation (2.22). In real world problems the permeability can vary by a factor of 10^9 or more on the domain, since for example the permeability of gravel $\sim 10^{-3}\text{cm}^2$, the permeability of sandstone $\sim 10^{-10}\text{cm}^2$ and the permeability of granite $\sim 10^{-16}\text{cm}^2$ (see Kuchling [86]). This can make the problem very ill-conditioned, as it will be shown in Chapter 4.
We then turned to the problem of finding the velocity $\mathbf{v} = -\alpha\nabla u$, rather than the pressure u. Therefore we considered the saddle point problem (2.30, 2.31) for which we derived the mixed weak formulation (2.34, 2.35) with solution $(\mathbf{v}, u) \in H_{0,N}(\text{div}, \Omega) \times L_2(\Omega)$. Its approximate formulation in finite dimensional subspaces $\mathcal{R}_h \subset H_{0,N}(\text{div}, \Omega)$ and $\mathcal{S}_h \subset L_2(\Omega)$ was given. Using this mixed formulation we derived the linear system (2.42), which is indefinite and therefore makes a direct application of any of the powerful classical methods for positive definite systems difficult. We therefore ideally would like to reduce this system to a positive definite system. Exactly this was done in Section 2.4.
For this reduction Raviart-Thomas elements were defined. We then decoupled the velocity from the pressure, which leads to system (2.45). Apart from being smaller than the original saddle point system

$$\begin{bmatrix} W(\alpha) & B \\ B^T & 0 \end{bmatrix} \begin{pmatrix} \mathbf{V} \\ \mathbf{U} \end{pmatrix} = \begin{pmatrix} \mathbf{G} \\ \mathbf{0} \end{pmatrix} \tag{2.49}$$

and symmetric positive definite, we can also apply all the preconditioners that will be developed in Chapters 4 to 6 easily to the new system

$$\mathring{A}(\alpha)\mathring{\mathbf{V}} = \begin{bmatrix} A(\alpha) & C(\alpha) \\ C^T(\alpha) & D(\alpha) \end{bmatrix} \mathring{\mathbf{V}} = \mathring{\mathbf{G}}, \tag{2.50}$$

and therefore can profit from the improved performance of our new domain decomposition method for groundwater flow problems. We will come back to this formulation of the problem in Section 7.5, where we will apply the domain decomposition preconditioners with different coarsening techniques to it.

Chapter 3

Random fields and maximum estimates

"*In a world as crazy as this one, it ought to be easy to find something that happens solely by chance. It isn't.*"

— Kevin McKeen, "The Orderly Pursuit of Pure Disorder", 1981 —

"*Arguments derived from probabilities are idle.*"

— Plato (Ancient Greek Philosopher, 427 BC-347 BC) —

3.1 Introduction

Random fields have been studied for over 100 years now. Elementary texts on general probability theory have been written by Ross [113], Hoel et al. [72], Grimmett and Welsh [62] and Stirzaker [124]. Higher level texts include Billingsley [11], Breiman [18], Kallenberg [82] and Shiryayev [119]. A good overview of general random fields was given in Vanmarcke [129], which also forms the basis of the following introduction. Gaussian random fields are discussed in more detail by Adler [4], Dudley [46], Hawkes [69], Lévy [88], Qualls and Watanabe [110] and Rozanov [114]. Publications on ideas about the distributions, etc. of maxima of Gaussian processes include Adler [3], Aldous [5], Kac and Slepian [81], Leadbetter et al. [87], Lindgren [90] and Pickands [108]. We will come back to this in more detail in Subsection 3.4.

In the analysis of the convergence rate of the unpreconditioned and preconditioned

systems for groundwater flow problems, we will have to deal repeatedly with *maxima of Gaussian random fields* (as they will be defined in Definition 3.2). It would therefore be of special interest to find good estimates for the expectation of the maxima of our fields or even better for the probability distributions of these maxima.

Using theoretical bounds for the expectation of the maxima of the underlying random fields, we will then be able to give estimates for the condition number and for the average number of conjugate gradient iterations of the unpreconditioned and domain decomposition preconditioned systems (see Chapters 4 and 5). Using the probability distributions (and the results from Appendix B) we will even be able to give estimates on how many cg-iterations it will take in 90%, 95%, etc. of the cases to obtain a certain accuracy of the solution of the system. We will therefore be able to give a full stochastic analysis of our preconditioner.

But let us start with the basics and first of all describe how the field α is determined. For simplicity consider *one single random variable* X with *probability density function* $f(X)$. The *expectation* of a function $g(X)$ is then defined by

$$\mathbb{E}\left[g(X)\right] := \int_{-\infty}^{+\infty} g(x)f(x)\mathrm{d}x, \tag{3.1}$$

provided this integral exists. (For the integral to exist, the product of the functions $g(X)$ and $f(X)$ must be "absolutely integrable", that is: $\mathbb{E}[|g(X)|]$ must be finite.)
If now $g(X) = X^n$, the n-th *moment* of X is defined as the expectation of g and the first moment is the *mean* or *expectation* of X

$$m := m_X := \mathbb{E}\left[X\right]. \tag{3.2}$$

The *variance* is then defined by

$$\sigma^2 := \sigma_X^2 := \mathrm{Var}\left[X\right] := \mathbb{E}\left[(X-m)^2\right] = \mathbb{E}\left[X^2\right] - m^2. \tag{3.3}$$

The expectation operation is linear and can therefore be interchanged with other linear operations such as summation, integration and differentiation.

After having recalled these standard definitions, remember that the probability density

function of a *Gaussian random variable* Z is

$$f(z) := \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{1}{2}\left(\frac{z-m}{\sigma}\right)^2\right\}, \quad \text{for } -\infty \leq z \leq +\infty. \tag{3.4}$$

and it is determined completely by its *mean* $m := \mathbb{E}[Z]$ and *variance* $\sigma^2 := \mathbb{E}[(Z - \mathbb{E}[Z])^2]$.

DEFINITION 3.1
Let Z be a Gaussian random variable. Then we call the random variable Y defined by

$$Y := \exp(Z), \tag{3.5}$$

a lognormal Gaussian random variable.

If Z is a Gaussian random variable with mean m and variance σ^2, the probability density function of $Y := \exp(Z)$ is

$$f_Y(y) = \frac{1}{y\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{\ln(y) - \ln(\tilde{m}_Y)}{\sigma}\right)^2\right\}, \quad y \geq 0, \tag{3.6}$$

where $\tilde{m}_Y$ denotes the median of Y, i.e. the value y_0 of Y, such that $\int_{-\infty}^{y_0} f_Y(y)dy = 0.5$. The *mean value* of the lognormal random variable $Y = e^Z$ is then given by (see Vanmarcke [129], p. 53ff.)

$$\mathbb{E}[Y] = \mathbb{E}[e^Z] = \exp\left\{m + \frac{1}{2}\sigma^2\right\}. \tag{3.7}$$

It has got the k-th moment

$$\mathbb{E}[Y^k] = \mathbb{E}[e^{kZ}] = \exp\left\{km + \frac{1}{2}k^2\sigma^2\right\} \tag{3.8}$$

and the variance

$$\sigma_Y^2 = \left(\exp\{\sigma^2\} - 1\right)\exp\{2m + \sigma^2\}. \tag{3.9}$$

Furthermore we know, for a function depending on *several random variables*, that the expectation of a function $g(Z_1, ..., Z_M)$ of several random variables $Z_1, ..., Z_M$ is

$$\mathbb{E}[g(Z_1, ..., Z_M)] = \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} g(z_1, ..., z_M) f(z_1, ..., z_M) \mathrm{d}z_1 ... \mathrm{d}z_M, \tag{3.10}$$

provided the multiple integral remains finite.

This can now be generalised to an infinite dimensional setting:

DEFINITION 3.2
Consider an open domain $\Omega \in \mathbb{R}^d$ and let each point $\mathbf{x} \in \Omega$ be associated with a random variable $Z(\mathbf{x})$. We then call Z a **random field** *on Ω. This field is said to be Gaussian, if and only if for each $n \in \mathbb{N}$ and each $\mathbf{x}_1, ..., \mathbf{x}_n \in \Omega$, the set of random variables $Z(\mathbf{x}_1), ..., Z(\mathbf{x}_n)$ is Gaussian. It can then be specified by its mean $m(\mathbf{x}) := \mathbb{E}[Z(\mathbf{x})]$ and covariance function,*

$$\Sigma(\mathbf{x}, \mathbf{y}) := \mathbb{E}[(Z(\mathbf{x}) - m(\mathbf{x}))(Z(\mathbf{y}) - m(\mathbf{y}))], \quad \mathbf{x}, \mathbf{y} \in \Omega. \tag{3.11}$$

The coefficient α in the second order partial differential equation (2.8, 2.9, 2.10) will be taken to be a *lognormal random field*, i.e.

$$\alpha(\mathbf{x}) := \exp\{Z(\mathbf{x})\}, \tag{3.12}$$

where $Z(\mathbf{x})$ is a Gaussian random field, with further properties defined as follows.
We are mainly interested in *statistically homogeneous isotropic* Gaussian random fields of Ornstein-Uhlenbeck type, which are defined as Gaussian random fields with constant mean $m(\mathbf{x}) \equiv 0$ for all $\mathbf{x} \in \Omega$ and having the covariance function

$$\Sigma(\mathbf{x}, \mathbf{y}) := \sigma^2 \exp\{-\|\mathbf{x} - \mathbf{y}\|_2 / \lambda\}, \tag{3.13}$$

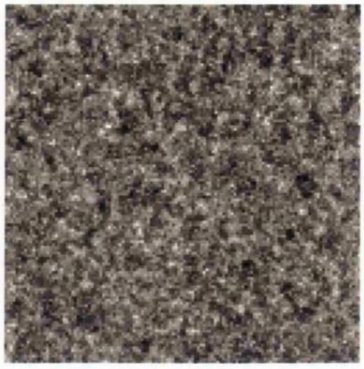
for positive constants σ and λ, where λ is called the *correlation length* and represents the length scale over which the field is correlated. Note that σ^2 is the variance of $Z(\mathbf{x})$ for all $\mathbf{x} \in \Omega$.
Expanding the covariance (3.11), we get

$$\Sigma(\mathbf{x}, \mathbf{y}) := \mathbb{E}[Z(\mathbf{x})Z(\mathbf{y})] - \mathbb{E}[Z(\mathbf{x})]\mathbb{E}[Z(\mathbf{y})]. \tag{3.14}$$

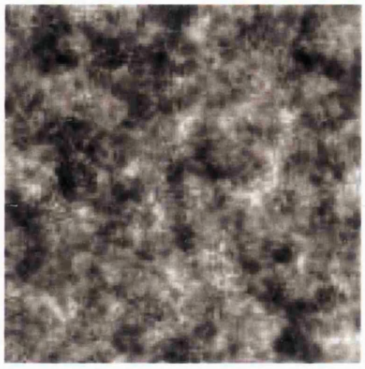
Thus, since $\mathbb{E}[Z(\mathbf{x})] = 0$ for all $\mathbf{x}$, and using (3.13), we have

$$\Sigma(\mathbf{x}, \mathbf{y}) := \mathbb{E}[Z(\mathbf{x})Z(\mathbf{y})] = \sigma^2 \exp\{-\|\mathbf{x} - \mathbf{y}\|_2 / \lambda\}. \tag{3.15}$$

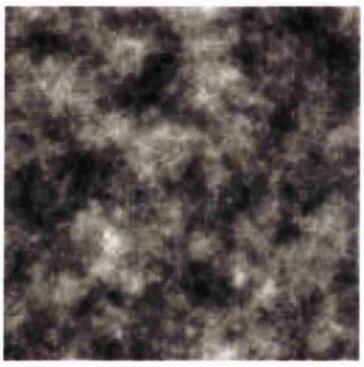
Increasing the correlation length λ therefore increases the "smoothness" of the field, i.e. the field changes less over a short distance as the following figures show, where the absolute values of the field depend on the variance:



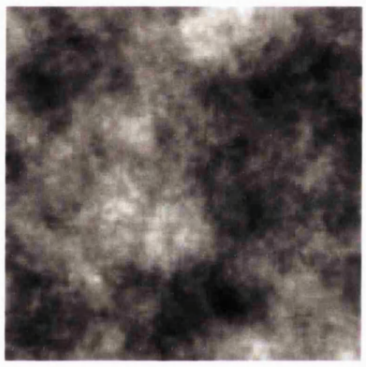
(a) Correlation length $\lambda = 1h$



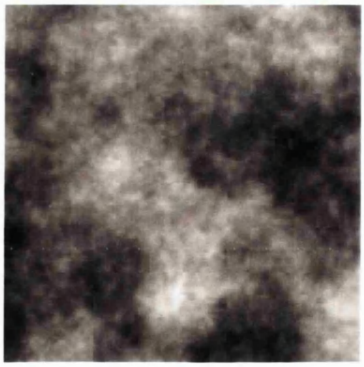
(b) Correlation length $\lambda = 5h$



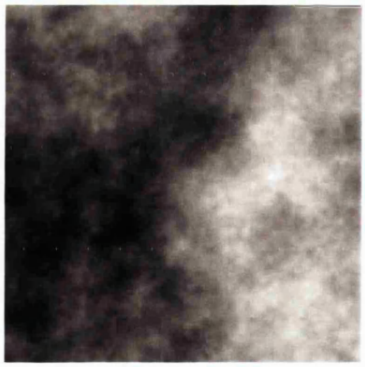
(c) Correlation length $\lambda = 10h$



(d) Correlation length $\lambda = 20h$



(e) Correlation length $\lambda = 50h$



(f) Correlation length $\lambda = 100h$

Figure 3-1: Dependence of the Gaussian random field on the correlation length λ (on the unit square and for $h = 1/128$)

Experiments show (see Dagan and Neumann [33]) that the *exponential* covariance function (3.13) describes fields occurring in groundwater problems well enough, so that (3.12) is a useful model for the permeability. Detailed discussions about the modeling of these fields and their use in groundwater flow models are given for example by Bras and Rodriguez-Iturbe [17], Ababou [2] and Harter [68].

One of the oldest and also hardest problems in the study of random fields is the determination of the probability

$$\mathbb{P}\left[\sup_{\mathbf{x}\in\Omega}\left(Z(\mathbf{x})\right) \leq b\right], \tag{3.16}$$

for a chosen value b and a random field Z, with a specific covariance function.
It is a fact, that, even today, there is no explicit formula for this simple probability in the general Gaussian situation, although such formulae exist for fields with some special covariance functions. However it is possible to find simple estimates for the general case without taking into consideration the specific covariance function of a field and also (significantly more precise) estimates for the isotropic Gaussian random fields with covariance of Ornstein-Uhlenbeck type, i.e. with covariance function

$$\Sigma\left(\mathbf{x},\mathbf{y}\right) := \sigma^2 \exp\left\{-\left\|\mathbf{y}-\mathbf{x}\right\|_2 / \lambda\right\}. \tag{3.17}$$

We will return to this topic after the following section.

3.2 Generation of random fields

After having defined the fields that will be used for the numerical experiments in the next chapters, it is now important to find ways to generate these fields as accurately as possible, but also in a computationally cheap way. The latter is significant, as the computations have to be performed on a large number of realisations of the field in the context of Monte Carlo methods.
In practice we replace the lognormal random field α on each element $\tau \in \mathcal{T}^h(\bar{\Omega})$ by the constant value $\alpha^\tau := \frac{1}{|\tau|}\int_\tau \alpha$.
With this simplification in mind several successful methods for the generation of zero-mean Gaussian random fields with given covariance function Σ have been proposed. Let $Z(\mathbf{x})$ again denote a random field with mean 0 and covariance function $\Sigma(\mathbf{x},\mathbf{y})$, $\mathbf{x},\mathbf{y} \in \Omega$.
Then fix a finite subset $X = \{\mathbf{x}_1,...,\mathbf{x}_n\} \subset \Omega$ of n points $\mathbf{x}_i$ for which we want to sample the Gaussian random field Z. We now want to compute a vector $\mathbf{Z} :=$

$(Z(\mathbf{x}_1), ..., Z(\mathbf{x}_n))^T$ of Gaussian random variables with mean 0 and covariance Σ at these points. Having made this choice, the following three methods among others have been very popular in the last few decades.

- **Cholesky decomposition method:** Let us introduce the covariance matrix C, which is defined in the following way:

  DEFINITION 3.3
  If $\mathbb{E}[\mathbf{Z}] = \mathbf{0}$, the covariance matrix C of $\mathbf{Z}$ is the symmetric $n \times n$ matrix with elements $C_{jk}(\mathbf{Z}) := \mathbb{E}[Z(\mathbf{x}_j) Z(\mathbf{x}_k)]$ (provided they all exist). In other words

  $$C(\mathbf{Z}) := \mathbb{E}\left[\mathbf{Z}^T\mathbf{Z}\right].$$

  This is equivalent to the evaluation of the covariance function Σ at pairs of points $\mathbf{x}_j$ and $\mathbf{x}_k$, $j, k = 1, ..., n$.
  (For arbitrary $\mathbf{Z}$ define $C(\mathbf{Z})$ to be the matrix $\mathbb{E}\left[(\mathbf{Z} - \mathbb{E}[\mathbf{Z}])^T (\mathbf{Z} - \mathbb{E}[\mathbf{Z}])\right]$.)

  By its definition it follows directly that $C(\mathbf{Z})$ is symmetric. Furthermore it was shown for example in Feller [56], Volume II, pages 81-82, that the covariance matrix of any non-degenerate probability distribution is positive definite.
  We can therefore find the Cholesky factorization of the covariance matrix $C(\mathbf{Z})$ and use it for the computation of $\mathbf{Z}$ (see Ripley [112], Section 4.5 and Cressie [32], Section 3.6) such that

  $$C = LL^T,$$

  where L is a lower triangular matrix. Knowing L, $\mathbf{Z}$ can now be generated by

  $$\mathbf{Z} = L\boldsymbol{\epsilon}, \tag{3.18}$$

  where $\boldsymbol{\epsilon} := (\epsilon_1, ..., \epsilon_n)^T$ is a vector of n independent and identically distributed (iid.) random variables ϵ_i in $N(0, 1)$, i.e. normally distributed with mean 0 and variance 1. It is easy to see that

  $$\mathbb{E}[\mathbf{Z}] = \mathbb{E}[L\boldsymbol{\epsilon}] = L\mathbb{E}[\boldsymbol{\epsilon}] = \mathbf{0},$$

  $$\mathrm{Var}[\mathbf{Z}] = \mathrm{Var}[L\boldsymbol{\epsilon}] = L\mathrm{Var}[\boldsymbol{\epsilon}] L^T = C.$$

  This method is simple and exact, but we have to compute the Cholesky factorization of a dense $n \times n$ matrix. With standard modern personal computers this is only possible for matrices of size up to approximately $n \times n \approx 10,000 \times 10,000$.

Incomplete Cholesky allows us to deal with larger problems, however the error analysis gets quite complex.

- **Turning bands method:** This algorithm involves the generation of a number of one-dimensional fields along lines radiating from a coordinate origin. These are then projected (as described in detail for example by Manatoglou and Wilson [96]) and combined at arbitrary points in space, which then leads to a sequence of discrete values or realisations of the field. With a clever choice of these lines, the method is computationally efficient by converting one dimensional problems to a higher dimensional field (see also Matheron [98], Brooker [21], Christakos [30], Manatoglou [95], Ripley [112], Tompson et al. [127] and Cressie [32]).

- **Circulant embedding:** This method was developed and analysed among others by Davis and Harte [35], Dembo et al. [36], Dietrich and Newsam [38], Harter [68], Wood and Chan [132] and Chan [26].
  Let us first of all consider the one-dimensional case and suppose, we need to find a realisation of an isotropic random field at n equally spaced points on the unit interval $[0,1]$, i.e. $\mathbf{Z} := (Z(0), Z(1/n), ..., Z((n-1)/n))^T$. Since the field is isotropic, the covariance Σ now only depends on the distance of two points. The covariance matrix is then given by

$$C = \begin{bmatrix} \Sigma(0) & \Sigma\left(\frac{1}{n}\right) & \dots & \Sigma\left(\frac{n-1}{n}\right) \\ \Sigma\left(\frac{1}{n}\right) & \Sigma(0) & \dots & \Sigma\left(\frac{n-2}{n}\right) \\ \vdots & \vdots & \ddots & \vdots \\ \Sigma\left(\frac{n-1}{n}\right) & \Sigma\left(\frac{n-2}{n}\right) & \dots & \Sigma(0) \end{bmatrix}. \tag{3.19}$$

The idea of the circulant embedding approach can in essence be described in the following two steps:

Random field generation by circulant embedding (1D):

(i) Choose $m := 2^g$ for some integer g, such that $2^g \geq 2(n-1)$. Embed $C(\mathbf{Z})$ in a $m \times m$ circulant covariance matrix Γ, i.e. a matrix of which the columns are circular shifts of the first column.

(ii) Use the fast Fourier transform twice in the way that we will describe below to create a random vector $\mathbf{Y} = (Y_0, Y_1, ..., Y_{m-1})^T$ of m variables Y_i which are iid. with mean 0 and covariance Σ, denoted by $N_m(0, C(\mathbf{Y}))$. We can then extract a vector $\mathbf{Z}$ of length n from this vector with the desired properties, i.e. being a sample of the Gaussian random field at points $\{0, 1/n, ..., (n-1)/n\}$.

Details of step (i):

Let

$$\gamma_j := \begin{cases} \Sigma\left(\frac{j}{n}\right), & 0 \leq j \leq \frac{m}{2}, \\ \Sigma\left(\frac{m-j}{n}\right), & \frac{m}{2} < j \leq m-1. \end{cases} \tag{3.20}$$

Then define the symmetric circulant matrix G by

$$G := \begin{bmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_{m-1} \\ \gamma_{m-1} & \gamma_0 & \cdots & \gamma_{m-2} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_1 & \gamma_2 & \cdots & \gamma_0 \end{bmatrix}. \tag{3.21}$$

Note that if $m \geq 2(n-1)$, the $n \times n$ top left hand submatrix of G is equal to C. For the construction of the vector $\mathbf{Z}$, we need G to be nonnegative definite. If this is not the case, increase g by 1. Repeat this until G is nonnegative definite (which, as one can prove, is the case for large enough g for strongly decaying covariance functions Σ).

Details of step (ii):

Now assume that we have chosen g large enough so that G is nonnegative definite. A standard result (see Brockwell and Davis [20]) shows that $G = Q\Lambda Q^*$, where $\Lambda := \operatorname{diag}\{\lambda_0, ..., \lambda_{m-1}\}$ is a diagonal matrix of eigenvalues of G, $Q :=$

$\{q_{jk} : 0 \leq j,k \leq m-1\}$ with

$$q_{j,k} := m^{-1/2} \exp\left(-\frac{2\pi ijk}{m}\right),$$

is a unitary Fourier matrix, Q^* its transpose and $i = \sqrt{-1}$.
Now let $\boldsymbol{\epsilon} := (\epsilon_0, ..., \epsilon_{m-1})^T$ be a vector of m iid. distributed values ϵ_i in $N(0,1)$ and define $\mathbf{Y} := Q\Lambda^{1/2}Q^*\boldsymbol{\epsilon}$. We can then show (similarly as for the Cholesky factorization method) that $\mathbf{Y}$ is $N_m(0, C(\mathbf{Y}))$. A subvector $\mathbf{Z}$ of $\mathbf{Y}$ of length n then is $N_n(0, C(\mathbf{Z}))$. The problem here however is, that m can be very large, i.e. in many cases of interest $m > 2^{15}$. Thus we have to find a way to compute $\mathbf{Y} := Q\Lambda^{1/2}Q^*\boldsymbol{\epsilon}$ quickly and with minimal storage. This can for example be done in the following way:

(i) Find the eigenvalues $\lambda_0, \lambda_1, ..., \lambda_{m-1}$ of G using a discrete Fourier transform of the sequence $\{\gamma_0, \gamma_1, ..., \gamma_{m-1}\}$, i.e.

$$\lambda_k = \sum_{j=0}^{m-1} \gamma_j \exp\left(-\frac{2\pi ijk}{m}\right), \quad k = 0, 1, ..., m-1. \tag{3.22}$$

(ii) Simulate $Q^*\boldsymbol{\epsilon}$ and determine (at the same time) $\mathbf{W} := \Lambda^{1/2}Q^*\boldsymbol{\epsilon}$.

(iii) Calculate $\mathbf{Z} := \{Z(0), Z(1/n), ..., Z((n-1)/n)\}^T$, where

$$Z\left(\frac{k}{n}\right) := \sum_{j=0}^{m-1} \left(W_j/m^{1/2}\right) \exp\left(-\frac{2\pi ijk}{m}\right), \quad k = 0, ..., n-1. \tag{3.23}$$

Essential here is, that (3.22) and (3.23) can be computed very efficiently using the fast Fourier transform (see for example Brockwell and Davis [20] and Numerical Analysis Group Limited [107]), provided that m is highly composite (e.g. if $m = 2^g$ for some integer g).

Harter [68] and Chan [26] also give similar algorithms for higher dimensional fields and vector-valued Gaussian random fields. They show that this method is computationally significantly cheaper than direct methods like the Cholesky factorization for large n and that it uses less storage than them.

We use the circulant embedding technique for our numerical experiments. The tests in 1D and 2D were performed with the "Gaussian" library, written by Boris Kozintsev

as part of his Ph.D. (see Kozintsev [84, 85]), which proved more accurate than other codes for our problem. For fields in 3D we use the "RandomField"-library in R (see Venables et al. [130]).

3.3 Maxima for general random fields

In the proofs that are given in this section a few theorems on random variables are of particular importance and we therefore would like to quote them in this report. They can also be found for example in (Adler [4], page 9).

THEOREM 3.4
(Hölder's inequality) *Let Z and Y be random variables, and p, q such that $p > 1$, $q > 1$, $\frac{1}{p} + \frac{1}{q} = 1$. Then, provided the right hand side is finite*

$$\mathbb{E}[|ZY|] \leq (\mathbb{E}[|Z|^p])^{1/p} (\mathbb{E}[|Y|^q])^{1/q}. \tag{3.24}$$

When $p = q = 2$ this is known as the Cauchy-Schwarz inequality.

THEOREM 3.5
(Jensen's inequality) *Let Z be a random variable and ψ a convex function on an interval $I \subseteq (-\infty, \infty)$. That is, for any $\lambda \in (0,1)$ and $u, v \in I$, $u < v$,*

$$\psi(u + \lambda(v - u)) \leq \psi(u) + \lambda(\psi(v) - \psi(u)). \tag{3.25}$$

Then, if the range of Z is in I and both $\mathbb{E}[|Z|]$ and $\mathbb{E}[|\psi(Z)|]$ are finite

$$\psi(\mathbb{E}[Z]) \leq \mathbb{E}[\psi(Z)]. \tag{3.26}$$

Recall that by 3.12, $\alpha(\mathbf{x}) := \exp\{Z(\mathbf{x})\}$. Let us approximate at the moment α by a piecewise constant function, which is given on each grid element τ by

$$\alpha^\tau := \alpha(\mathbf{m}^\tau) = \exp\{Z(\mathbf{m}^\tau)\}, \tag{3.27}$$

where $\mathbf{m}^\tau$ is the centroid of τ.
This simplification is reasonable, if we choose the triangulation fine enough. Let us now discuss how to find maxima and minima of such a random field α without taking

its covariance into consideration. We therefore assume that the uniform triangulation $\mathcal{T}^h(\bar{\Omega})$ is of size h, and that $\#\mathcal{T}_\Omega^h$ is the number of elements τ in $\bar{\Omega}$. Our aim now is to find bounds for the expectation of the maximum of the vector $\boldsymbol{\alpha}$, which we define to consist of the values α^τ for all $\tau \in \mathcal{T}^h(\bar{\Omega})$.

In the following lemmata of this subsection we always assume $Z(\mathbf{x})$, $\mathbf{x} \in \Omega$, to be Gaussian with mean 0 independent of $\mathbf{x}$ and we consider fields α without considering their specific covariance function.

We now would like to find a general bound for the maximum value of such a vector. The following bounds are new results and based on simple norm inequalities. They work for general random fields, but can therefore be quite pessimistic for random fields with strong covariance.

LEMMA 3.6
For any $p \geq 1$ we have $\mathbb{E}[\max_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)] \leq \exp\{\frac{1}{2}p\sigma^2\}\{\#\mathcal{T}_\Omega^h\}^{1/p}$.

Proof:
Since α^τ is a positive valued random variable, we get

$$\mathbb{E}[\max_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)] = \mathbb{E}[||\boldsymbol{\alpha}||_\infty] \leq \mathbb{E}[||\boldsymbol{\alpha}||_p] = \mathbb{E}\left[\{\sum_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)^p\}^{1/p}\right],$$

where we have used the vector norm inequality $||\mathbf{x}||_\infty \leq ||\mathbf{x}||_p$ for an n-dimensional vector $\mathbf{x}$ and any p-norm.
If $p = 1$, the result follows directly, as then

$$\mathbb{E}[\max_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)] = \mathbb{E}\left[\sum_{\tau\in\mathcal{T}^h(\bar{\Omega})}\alpha^\tau\right] = \{\#\mathcal{T}_\Omega^h\}\cdot\mathbb{E}[\alpha^\tau] = \{\#\mathcal{T}_\Omega^h\}\cdot\exp\{\frac{1}{2}\sigma^2\}.$$

Otherwise, using Hölder's inequality (see Theorem 3.4) for $p > 1$:

$$\mathbb{E}[\max_{\tau\in\mathcal{T}^h(\bar{\Omega})}\alpha^\tau] \leq \left\{\mathbb{E}\left[\sum_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)^p\right]\right\}^{1/p}\{\mathbb{E}[1]\}^{1/q} = \{\sum_{\tau\in\mathcal{T}^h(\bar{\Omega})}\mathbb{E}[(\alpha^\tau)^p]\}^{1/p}.$$

Now observe that

$$(\alpha^\tau)^p = \exp\{p\cdot Z(\mathbf{m}^\tau)\},$$

and note that $p \cdot Z(\mathbf{m}^\tau)$ is a normal random variable with mean 0 and variance $p^2\sigma^2$. Therefore an application of (3.7) implies that

$$\mathbb{E}[(\alpha^\tau)^p] = \exp\{\frac{1}{2}p^2\sigma^2\},$$

and so it follows that

$$\begin{aligned}\mathbb{E}[\max_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)] &\leq \{\sum_{\tau\in\mathcal{T}^h(\bar{\Omega})}\mathbb{E}[(\alpha^\tau)^p]\}^{1/p} = \{\sum_{\tau\in\mathcal{T}^h(\bar{\Omega})}\exp\{\frac{1}{2}p^2\sigma^2\}\}^{1/p} \\ &= \exp\{\frac{1}{2}p\sigma^2\}\cdot\{\#\mathcal{T}_\Omega^h\}^{1/p}.\end{aligned} \tag{3.28}$$

□

The best choice of p is found in the following lemma.

LEMMA 3.7

$$\mathbb{E}[\max_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)] \leq \begin{cases}\exp\{\sqrt{2}\sigma\sqrt{\ln(\#\mathcal{T}_\Omega^h)}\}, & \text{when } \ln(\#\mathcal{T}_\Omega^h) \geq \frac{\sigma^2}{2}, \\ \exp\{\frac{1}{2}\sigma^2\}\{\#\mathcal{T}_\Omega^h\}, & \text{otherwise.}\end{cases} \tag{3.29}$$

Proof:
As $\mathbb{E}[\max_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)] \leq \exp\{\frac{1}{2}p\sigma^2\}\{\#\mathcal{T}_\Omega^h\}^{1/p}$, we therefore try to minimise the function $f(p) := \exp\{\frac{1}{2}p\sigma^2\}\{\#\mathcal{T}_\Omega^h\}^{1/p}$ for $p > 1$.
Differentiating yields

$$f'(p) = \exp\{\frac{1}{2}p\sigma^2\}\left(-\frac{1}{p^2}\right)\{\#\mathcal{T}_\Omega^h\}^{1/p}\ln(\#\mathcal{T}_\Omega^h) + \frac{1}{2}\sigma^2\exp\{\frac{1}{2}p\sigma^2\}\{\#\mathcal{T}_\Omega^h\}^{1/p}.$$

From this we see that $f'(p) = 0$ if and only if

$$-\frac{1}{p^2}\cdot\ln(\#\mathcal{T}_\Omega^h) + \frac{1}{2}\sigma^2 = 0, \tag{3.30}$$

which is equivalent to

$$p^2 = \frac{2}{\sigma^2}\cdot\ln(\#\mathcal{T}_\Omega^h). \tag{3.31}$$

We therefore get the *local extremum* for

$$p^* = \frac{\sqrt{2}}{\sigma}\sqrt{\ln(\#\mathcal{T}_\Omega^h)}. \tag{3.32}$$

As $f''(p^*) > 0$, this is a *local minimum* of $f(p)$ over $p \in \mathbb{R}$. However in this application, p is constrained to satisfy $p \geq 1$. If $p^* \geq 1$ then, since $f(p) \to \infty$ for $p \to \infty$, p^* is a *global minimiser* of $f(p)$ over $p \in [1, \infty]$. When $p^* < 1$, we have

$$f(1) = \min_{p\in[1,\infty]} f(p), \tag{3.33}$$

and the result follows. □

As we will see later, we do not only need estimates for the maxima of the entries of such vectors $\boldsymbol{\alpha}$, but also for their minima. A general bound can be found in a similar way, using the following Lemma.

LEMMA 3.8

$$\mathbb{E}[\min_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)] \geq \exp\{-\frac{1}{2}p\sigma^2\}\{\#\mathcal{T}_\Omega^h\}^{-1/p} \text{ for any } p \geq 1. \tag{3.34}$$

Proof:

First we write

$$\mathbb{E}\left[\frac{1}{\min_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)}\right] = \mathbb{E}\left[\left\|\boldsymbol{\alpha}^{-1}\right\|_\infty\right], \qquad \text{where } \boldsymbol{\alpha}^{-1} := \left(\frac{1}{\alpha^\tau}\right)_{\tau\in\mathcal{T}^h(\bar{\Omega})} \tag{3.35}$$

Thus, as in Lemma 3.6,

$$\mathbb{E}\left[\frac{1}{\min_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)}\right] \leq \mathbb{E}\left[\left\|\boldsymbol{\alpha}^{-1}\right\|_p\right] = \mathbb{E}\left[\left(\sum_{\tau\in\mathcal{T}^h(\bar{\Omega})}\left(\frac{1}{\alpha^\tau}\right)^p\right)^{1/p}\right]. \tag{3.36}$$

When $p > 1$, we can use Hölder's inequality (see Theorem 3.4) to obtain

$$\mathbb{E}\left[\frac{1}{\min_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)}\right] \leq \left(\mathbb{E}\left[\sum_{\tau\in\mathcal{T}^h(\bar{\Omega})}\left(\frac{1}{\alpha^\tau}\right)^p\right]\right)^{1/p}\left(\mathbb{E}[1]^{1/q}\right), \tag{3.37}$$

where $\frac{1}{p} + \frac{1}{q} = 1$.
Hence

$$\mathbb{E}\left[\frac{1}{\min_{\tau \in \mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)}\right] \leq \left(\sum_{\tau \in \mathcal{T}^h(\bar{\Omega})} \mathbb{E}\left[\left(\frac{1}{\alpha^\tau}\right)^p\right]\right)^{1/p}, \tag{3.38}$$

For $p = 1$, (3.37) follows directly from (3.36).
Now recalling (3.27), we see that

$$\left(\frac{1}{\alpha^\tau}\right)^p = \exp\{-p \cdot Z(\mathbf{m}^\tau)\}.$$

Since $-p \cdot Z(\mathbf{m}^\tau)$ is a normal random variable with mean 0 and variance $p^2\sigma^2$, we get from (3.7)

$$\mathbb{E}\left[\left(\frac{1}{\alpha^\tau}\right)^p\right] = \exp\{\frac{1}{2}p^2\sigma^2\},$$

and so, from (3.38),

$$\begin{aligned}\mathbb{E}\left[\frac{1}{\min_{\tau \in \mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)}\right] &\leq \{\sum_{\tau \in \mathcal{T}^h(\bar{\Omega})} \mathbb{E}\left[\left(\frac{1}{\alpha^\tau}\right)^p\right]\}^{1/p} \\ &= \{\sum_{\tau \in \mathcal{T}^h(\bar{\Omega})} \exp\{\frac{1}{2}p^2\sigma^2\}\}^{1/p} = \exp\{\frac{1}{2}p\sigma^2\} \cdot \{\#\mathcal{T}_\Omega^h\}^{1/p}.\end{aligned} \tag{3.39}$$

Now recall that we are interested in a lower bound for $\mathbb{E}\left[\min_{\tau \in \mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)\right]$. To obtain this recall Jensen's inequality (Theorem 3.5).

Since $\phi(x) := 1/x$ is convex on $(0, \infty)$, we get

$$(\mathbb{E}[Z])^{-1} \leq \mathbb{E}\left[Z^{-1}\right]. \tag{3.40}$$

Put $Z := \min_{\tau \in \mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)$ and this yields

$$\mathbb{E}[\min_{\tau \in \mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)] \geq \left[\mathbb{E}\left[\frac{1}{\min_{\tau \in \mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)}\right]\right]^{-1} \geq \exp\{-\frac{1}{2}p\sigma^2\}\{\#\mathcal{T}_\Omega^h\}^{-1/p}. \tag{3.41}$$

□

LEMMA 3.9

$$\mathbb{E}[\min_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)] \geq \begin{cases} \exp\{-\sqrt{2}\sigma\sqrt{\ln(\#\mathcal{T}_\Omega^h)}\}, & \text{when } \ln(\#\mathcal{T}_\Omega^h) \geq \frac{\sigma^2}{2}, \\ \exp\{-\frac{1}{2}\sigma^2\}(\#\mathcal{T}_\Omega^h)^{-1}, & \text{otherwise.} \end{cases} \tag{3.42}$$

Proof:
We now have to maximise the function $g(p) := \exp\{-\frac{1}{2}p\sigma^2\}\{\#\mathcal{T}_\Omega^h\}^{-1/p}$, where $p > 1$. Let $f(p) := \exp\{\frac{1}{2}p\sigma^2\}\{\#\mathcal{T}_\Omega^h\}^{1/p}$.
Clearly $g(p) = 0$ if and only if $f'(-p) = 0$ and from Lemma 3.7 this holds (for $p \geq 1$) if and only if $p = p^* = (\sqrt{2}/\sigma)\sqrt{\ln(\#\mathcal{T}_\Omega^h)}$. The result then follows as in Lemma 3.7.
□

In the next chapter we will see that the condition number of the stiffness matrices that we consider depends linearly on $\left\{\max_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)\right\} / \left\{\min_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)\right\}$.
Using a Cauchy-Schwarz inequality we obtain

$$\mathbb{E}\left[\frac{\max_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)}{\min_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)}\right] \leq \mathbb{E}\left[\left(\max_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)\right)^2\right]^{1/2} \cdot \mathbb{E}\left[\left(\min_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)\right)^{-2}\right]^{1/2}. \tag{3.43}$$

Bounds for the right hand side of (3.43) are obtained in the next Lemma.

LEMMA 3.10
We have

$$\mathbb{E}[(\max_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau))^2] \leq \exp\{2p\sigma^2\}\{\#\mathcal{T}_\Omega^h\}^{1/p}, \tag{3.44}$$

$$\mathbb{E}[(\min_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau))^{-2}] \leq \exp\{2p\sigma^2\}\{\#\mathcal{T}_\Omega^h\}^{1/p} \tag{3.45}$$

and therefore

$$\mathbb{E}\left[\frac{\max_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)}{\min_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)}\right] \leq \exp\{2p\sigma^2\}\{\#\mathcal{T}_\Omega^h\}^{1/p} \tag{3.46}$$

for any $1 \leq p < \infty$.

Proof:
Using the fact, that $\alpha^\tau > 0$ is a positive valued random variable for all $\tau \in \mathcal{T}^h(\bar{\Omega})$, we

get:

$$\mathbb{E}[(\max_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau))^2] = \mathbb{E}[||\boldsymbol{\alpha}||_\infty^2] = \mathbb{E}[||\boldsymbol{\alpha}^2||_\infty] \le \mathbb{E}[||\boldsymbol{\alpha}^2||_p] = \mathbb{E}\left[\{\sum_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)^{2p}\}^{1/p}\right]. \tag{3.47}$$

For $p > 1$ it follows, from Hölder's inequality (Theorem 3.4), that

$$\mathbb{E}[(\max_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau))^2] \le \left(\mathbb{E}\left[\sum_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)^{2p}\right]\right)^{1/p}\left(\mathbb{E}[1]^{1/q}\right) = \left(\sum_{\tau\in\mathcal{T}^h(\bar{\Omega})}\mathbb{E}[(\alpha^\tau)^{2p}]\right)^{1/p}.$$

Using (3.47) directly for $p = 1$, we have

$$\mathbb{E}[(\max_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau))^2] \le \left(\sum_{\tau\in\mathcal{T}^h(\bar{\Omega})}\mathbb{E}[(\alpha^\tau)^{2p}]\right)^{1/p}, \tag{3.48}$$

for all $1 \le p < \infty$.

Now by (3.27), we get

$$(\alpha^\tau)^{2p} = \exp\{2p \cdot Z(\mathbf{m}^\tau)\},$$

and $2p \cdot Z(\mathbf{m}^\tau)$ is a normal random variable with mean 0 and variance $4p^2\sigma^2$. Therefore, again by (3.7), $\mathbb{E}[(\alpha^\tau)^{2p}] = \exp\{2p^2\sigma^2\}$ and (3.48) implies

$$\begin{aligned}\mathbb{E}[(\max_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau))^2] &\le \{\sum_{\tau\in\mathcal{T}^h(\bar{\Omega})}\mathbb{E}[(a^\tau)^{2p}]\}^{1/p} \\ &= \{\sum_{\tau\in\mathcal{T}^h(\bar{\Omega})}\exp\{2p^2\sigma^2\}\}^{1/p} = \exp\{2p\sigma^2\}\cdot\{\#\mathcal{T}_\Omega^h\}^{1/p}.\end{aligned} \tag{3.49}$$

In an equivalent way, we show (see (3.39)):

$$\mathbb{E}[(\min_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau))^{-2}] \le \exp\{2p\sigma^2\}\{\#\mathcal{T}_\Omega^h\}^{1/p}. \tag{3.50}$$

Therefore (using (3.43))

$$\begin{aligned}\mathbb{E}\left[\frac{\max_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)}{\min_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)}\right] &\le \mathbb{E}\left[\left(\max_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)\right)^2\right]^{1/2}\cdot\mathbb{E}\left[\left(\min_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)\right)^{-2}\right]^{1/2} \\ &\le \exp\{2p\sigma^2\}\{\#\mathcal{T}_\Omega^h\}^{1/p}.\end{aligned} \tag{3.51}$$

□

LEMMA 3.11

$$\mathbb{E}\left[\frac{\max_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)}{\min_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)}\right] \leq \begin{cases} \exp\{2\sqrt{2}\sigma\sqrt{\ln\left(\#\mathcal{T}_\Omega^h\right)}\}, & \text{for } \ln\left(\#\mathcal{T}_\Omega^h\right) \geq 2/\sigma^2, \\ \exp\{2\sigma^2\}\{\#\mathcal{T}_\Omega^h\}, & \text{otherwise.} \end{cases} \tag{3.52}$$

Proof:
Here we have to minimise the function $h(p) := \exp\{2p\sigma^2\}\left(\#\mathcal{T}_\Omega^h\right)^{1/p}$, where $p \geq 1$.
In a very similar way as in Lemma 3.7 and in Lemma 3.9 inequality (3.52) follows. □

3.4 Maxima for Gaussian random fields with covariance of Ornstein-Uhlenbeck type

As the numerical results in the next section will show the results of the previous section like Lemma 3.6, 3.8 and 3.10 (based mainly on norm inequalities) can be very pessimistic for fields with strong covariance. Despite long research in estimates for maxima of fields of this type, there is no closed formula for fields for domains of dimension $d > 1$ for neither the expectation nor the probability distribution of the maximum value. However David Aldous (see Aldous [5]) developed estimates that deliver very good estimates for probability distributions of the maxima of fields with certain covariance functions (as also shown in the next section).
We here want to focus on the heuristics developed for fields with covariance of Ornstein-Uhlenbeck type and want to apply them to our problem.

The heuristics are based on the study of random mechanisms governing the positioning of random sets in a domain Ω, so called *covering processes* (see Hall [67]). These sets are then identified as the areas of the random field, for which the field values exceed some (large) value b. For large enough b these regions become small islands (see Figure 3-2). Their size will depend, among other things, on the correlation length λ.

For details on Poisson clumping heuristics we refer the reader to Aldous [5]. We only want to give the results that are of importance for our estimates. Let us therefore first of all define the following.

DEFINITION 3.12
Consider a unit volume (i.e. a unit interval in 1D, a unit square in 2D, resp. a unit cube in 3D). The ***Poisson distribution*** *assigns probabilities to the number of points of a set of s identically and independently distributed points that lie in a fixed subset $\mathcal{S}$ of this unit volume. Its probability density function is given by*

$$f(j,\mu) = \frac{\exp\{-\mu\,|\mathcal{S}|\}\,(\mu\,|\mathcal{S}|)^j}{j!} \text{ for } j = 0,1,2,\ldots,s, \tag{3.53}$$

where μ is a parameter indicating the average number of points in the unit volume, the so called ***intensity*** *or* ***process rate****, and where $|\mathcal{S}|$ denotes the Lebesgue measure of $\mathcal{S}$ (as defined in Dudley [47], Subsection 3.5). In other words, this probability density function gives us the probability, that j of the s points lie in $\mathcal{S}$.*

DEFINITION 3.13
Let $\mathcal{C} \subset \mathbb{R}^d$ be a (random) set, which we obtain by picking sets from a list $\mathcal{B}_1, ..., \mathcal{B}_k$ with probabilities $p_1, ..., p_k$. A ***mosaic*** *$\mathcal{S}$ is then constructed by the following three step process:*

(i) Set down points $\mathbf{y}$ according to a Poisson point process with process rate μ per unit volume.

(ii) For each $\mathbf{y}$ pick a random subset $\mathcal{C}_\mathbf{y}$ distributed as $\mathcal{C}$, independent for different $\mathbf{x}$.

(iii) Finally set $\mathcal{S} := \bigcup(\mathbf{y} + \mathcal{C}_\mathbf{y})$, where $\mathbf{y} + \mathcal{C}_\mathbf{y} := \{\mathbf{y} + \mathbf{c} : \mathbf{c} \in \mathcal{C}_\mathbf{y}\}$

Call the $\mathbf{y}$'s the ***centers*** *and the sets $\mathbf{y} + \mathcal{C}_\mathbf{y}$ the* ***clumps*** *of the mosaic.*
A mosaic $\mathcal{S}$ is called ***sparse****, if it only covers a "small" portion of $\mathbb{R}^d$ in the sense that*

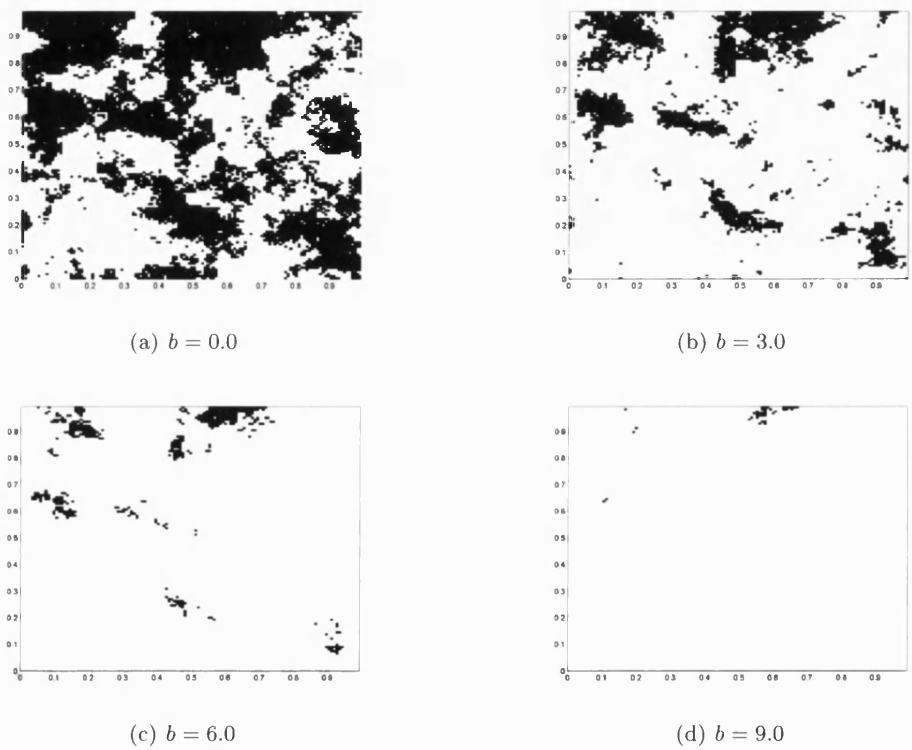
$$p := \mathbb{P}[\mathbf{x} \in \mathcal{S}], \text{ for } \mathbf{x} \in \mathbb{R}^d \text{ fixed},$$

is "small".

Having defined Poisson distributions and (sparse) mosaics, we now would like to focus on the supremum

$$M_\Omega := \sup_{\mathbf{x}\in\Omega} Z(\mathbf{x}) \tag{3.54}$$

of a random field Z on a domain $\Omega \subset \mathbb{R}^d$.
Fix therefore some "high" level b, where "high" means that only a small proportion of the field values is larger than b. (The numerical results in the next section show, that the heuristics do in fact a very good job, if 90% or more of the field values are smaller than b.)
Having fixed b consider the random set $\mathcal{S}_b := \{\mathbf{x} : Z(\mathbf{x}) \geq b\}$ next. Assume that this set resembles a mosaic process with some clump rate μ. Figure 3-2 displays $\mathcal{S}_b$ for various levels b.



(a) $b = 0.0$ (b) $b = 3.0$

(c) $b = 6.0$ (d) $b = 9.0$

Figure 3-2: Clumps for several values of b for a field α of Ornstein-Uhlenbeck type with $\lambda = 10h$ and $\sigma^2 = 8$ (Plots of $\log(\alpha)$, where $\mathcal{S}_b$ is plotted in black)

Denote $|\mathcal{S}_b|$ the volume of $\mathcal{S}_b$. Then it was shown in Aldous [5], Section J.5, that

$$\mathbb{P}[M_\Omega \leq b] \approx \exp\{-\mu_b |\Omega|\}, \tag{3.55}$$

where

$$\mu_b := \frac{\mathbb{P}[Z(\mathbf{x}) \geq b]}{\mathbb{E}[|\mathcal{S}_b|]} \tag{3.56}$$

(since by the definition of the clump rate $\mu_b \mathbb{E}[|\mathcal{S}_b|] = \mathbb{P}[\mathbf{x} \in \mathcal{S}_b] = \mathbb{P}[Z(\mathbf{x}) \geq b]$).
As in Aldous [5] "$\approx$" here means "is approximately equal" in a heuristic, rather vague sense. To any heuristic conclusion like (3.55) there corresponds a formal assertion as given in more detail in Aldous [5]. Similarly "$\lesssim$" will denote in the following "approximately less or equal" in this heuristic sense.
We now have to find a way to compute μ_b and therefore $\mathbb{E}[|\mathcal{S}_b|]$.

There exists no formula for μ_b for general covariance. For isotropic Gaussian processes however it was shown by Aldous [5], Section J18:

Let

$$f_0(b) := (2\pi)^{-d/2}\sigma^{-1}\exp\left\{-\frac{1}{2}db^2/\sigma^2\right\}, \qquad b \in \mathbb{R}, \tag{3.57}$$

i.e. the (multivariate, d-dimensional) Gaussian probability density function (with mean 0, variance σ^2 and no correlation) for a vector $\mathbf{b} := (b, ..., b)^T \in \mathbb{R}^d$ (compare (3.4)). To obtain numerical values for $\mathbb{P}[M_\Omega \leq b]$ (as defined in (3.54)) for isotropic random fields it is shown by Aldous [5], pages 205-207, that if $0 < q \leq 2$ and $0 < r < \infty$ are parameters and $Z(\mathbf{x})$, with $\mathbf{x} \in \mathbb{R}^d$ is a stationary mean-zero Gaussian random field with covariance of the form

$$\mathbb{E}[Z(\mathbf{x})Z(\mathbf{y})] := 1 - r\|\mathbf{y} - \mathbf{x}\|_2^q \text{ as } \|\mathbf{y} - \mathbf{x}\|_2 \to 0,$$

the clump rate is given by

$$\mu_b = K_{d,q} r^{d/q} b^{2d/q-1} f_0(b), \text{ for } b \text{ large}, \tag{3.58}$$

where f_0 is defined as in (3.57), $0 < K_{d,q} < \infty$ only depends on the dimension d of the domain and the parameter q.

In our case, with covariance of Ornstein-Uhlenbeck type, we get

$$\mathbb{E}[Z(\mathbf{x})Z(\mathbf{y})] := \exp\{-\|\mathbf{y} - \mathbf{x}\|_2/\lambda\} \approx 1 - \|\mathbf{y} - \mathbf{x}\|_2/\lambda \text{ as } \|\mathbf{y} - \mathbf{x}\|_2 \to 0,$$

and therefore $q \approx 1$ and $r \approx 1/\lambda$ (where λ again is the correlation length of the random field).

As shown by Aldous [5],

$$K_{d,2} = \pi^{-\frac{1}{2}d}, \text{ for d} \geq 1 \text{ and } K_{1,1} = 1. \tag{3.59}$$

Sadly this are the only values for which $K_{d,q}$ is known explicitly.
However this means that in our one dimensional example we get $K_{1,1} = 1$.

So for all other values of d and q, only bounds for μ_b are known. Those can be found for example using *Slepians inequality* (as discussed in some detail by Leadbetter et al. [87]), which we would like to quote here without proof:

LEMMA 3.14
Let Z and $\hat{Z}$ be Gaussian processes with mean zero and same variance. Suppose there exists a constant $\delta > 0$ such that $\mathbb{E}[Z(\mathbf{x})Z(\mathbf{y})] \leq \mathbb{E}\left[\hat{Z}(\mathbf{x})\hat{Z}(\mathbf{y})\right]$ for all $\|\mathbf{y} - \mathbf{x}\|_2 \leq \delta$. Then

$$\mathbb{P}\left[\sup_{\mathbf{x}\in\Omega} Z(\mathbf{x}) \geq b\right] \geq \mathbb{P}\left[\sup_{\mathbf{x}\in\Omega} \hat{Z}(\mathbf{x}) \geq b\right] \text{ for all } b, \text{ and } \Omega \text{ with } diam(\Omega) \leq \delta. \tag{3.60}$$

Using this lemma we can then prove the following lemma.

LEMMA 3.15
For a lognormal field α on $\Omega \in \mathbb{R}^d$ and with f_0 as defined in (3.57), we obtain for $d = 1$

$$\begin{aligned} \mathbb{P}\left[\sup_{\mathbf{x}\in\Omega}(\alpha(\mathbf{x})) \leq \exp\{b\}\right] &= \mathbb{P}\left[\inf_{\mathbf{x}\in\Omega}(\alpha(\mathbf{x})) \geq \exp\{-b\}\right] \\ &\approx \exp\left\{-f_0(b)\cdot\lambda^{-1}b\,|\Omega|\right\}. \end{aligned} \tag{3.61}$$

and for $d = 2$

$$\begin{aligned} \exp\left\{-f_0(b)\cdot\lambda^{-2}b^3\,|\Omega|\right\} &\lesssim \mathbb{P}\left[\sup_{\mathbf{x}\in\Omega}(\alpha(\mathbf{x})) \leq \exp\{b\}\right] \\ &= \mathbb{P}\left[\inf_{\mathbf{x}\in\Omega}(\alpha(\mathbf{x})) \geq \exp\{-b\}\right] \\ &\lesssim \exp\left\{-f_0(b)\cdot\lambda^{-2}b^3\cdot\frac{|\Omega|}{2}\cdot\right\}, \end{aligned} \tag{3.62}$$

where $f_0(b)$ is defined in (3.57).

Proof:
Let $\mu_{b,Z}$ denote the clump rate of Z and $\mu_{b,\hat{Z}}$ denote the clump rate of $\hat{Z}$. Then as

$\mu_{b,Z} = \frac{\mathbb{P}[Z(\mathbf{x})\geq b]}{\mathbb{E}[|\mathcal{S}_b|]}$ and $\mu_{b,\hat{Z}} = \frac{\mathbb{P}[\hat{Z}(\mathbf{x})\geq b]}{\mathbb{E}[|\mathcal{S}_b|]}$, inequality (3.60) implies

$$\mu_{b,Z} \geq \mu_{b,\hat{Z}}. \tag{3.63}$$

Now let $\mathbf{x} =: (x_1, \ldots, x_d), \mathbf{y} =: (y_1, \ldots, y_d) \in \mathbb{R}^d$ and $Z_r(\mathbf{x})$ be the field with covariance

$$\mathbb{E}[Z_r(\mathbf{x})Z_r(\mathbf{y})] := 1 - r\left(\sum_{i=1}^{d} |y_i - x_i|^q\right). \tag{3.64}$$

As shown in Aldous [5], the clump rate of this particular sparse mosaic is given by

$$\mu_{b,Z_r} = (K_{1,q})^d r^{d/q} b^{2d/q-1} f_0(b), \tag{3.65}$$

where f_0 is defined as in (3.57).
Remember that $\mathbb{E}[Z_1(\mathbf{0})Z_1(\mathbf{t})] := 1 - \|\mathbf{t}\|_q^q$, that $\mathbb{E}[Z(\mathbf{0})Z(\mathbf{t})] := 1 - \|\mathbf{t}\|_2^q$ and that $\mathbb{E}[Z_{d^{-1/2}}(\mathbf{0})Z_{d^{-1/2}}(\mathbf{t})] := 1 - d^{-1/2}\|\mathbf{t}\|_q^q$. Furthermore we know by simple norm inequalities

$$\|\mathbf{t}\|_q \geq \|\mathbf{t}\|_2 \geq d^{-1/(2q)}\|\mathbf{t}\|_q \text{ for } 0 < q \leq 2. \tag{3.66}$$

Using these inequalities we get

$$\mathbb{E}[Z_1(\mathbf{0})Z_1(\mathbf{t})] \leq \mathbb{E}[Z(\mathbf{0})Z(\mathbf{t})] \leq \mathbb{E}[Z_{d^{-1/2}}(\mathbf{0})Z_{d^{-1/2}}(\mathbf{t})]. \tag{3.67}$$

Slepian's inequality combined with (3.56) and (3.65) then leads to

$$d^{-\frac{d}{2q}}(K_{1,q})^d \leq K_{d,q} \leq (K_{1,q})^d, \tag{3.68}$$

and in particular for $q = 1$

$$d^{-\frac{d}{2}} \leq K_{d,1} \leq 1. \tag{3.69}$$

This gives us upper and lower bounds for the parameters $K_{d,q}$ and therefore for μ_b and the probability distribution of the maxima.

Now let $M_\Omega := \sup_{\mathbf{x}\in\Omega}(Z(\mathbf{x}))$, $m_\Omega := \inf_{\mathbf{x}\in\Omega}(Z(\mathbf{x}))$, $\hat{M}_\Omega := \sup_{\mathbf{x}\in\Omega}(\alpha(\mathbf{x}))$ and $\hat{m}_\Omega := \inf_{\mathbf{x}\in\Omega}(\alpha(\mathbf{x}))$.

Then we know since Z is a Gaussian random field with zero mean, that

$$\mathbb{P}[M_\Omega \leq b] = \mathbb{P}[m_\Omega \geq -b]. \tag{3.70}$$

Furthermore, we know since $\alpha(\mathbf{x}) := \exp\{Z(\mathbf{x})\}$, that

$$\mathbb{P}[M_\Omega \leq b] = \mathbb{P}\left[\hat{M}_\Omega \leq \exp\{b\}\right] \tag{3.71}$$

and

$$\mathbb{P}[m_\Omega \geq -b] = \mathbb{P}[\hat{m}_\Omega \geq \exp\{-b\}], \tag{3.72}$$

i.e. combining (3.70), (3.71) and (3.72)

$$\mathbb{P}[M_\Omega \leq b] = \mathbb{P}[m_\Omega \geq -b] = \mathbb{P}\left[\hat{M}_\Omega \leq \exp\{b\}\right] = \mathbb{P}[\hat{m}_\Omega \geq \exp\{-b\}]. \tag{3.73}$$

Now denote again $f_0(b) := (2\pi)^{-d/2}\sigma^{-1}\exp\{-db^2/(2\sigma^2)\}$.
By (3.58) we know $\mu_b = K_{d,q}r^{d/q}b^{2d/q-1}f_0(b)$, for b large, where for lognormal random fields with covariance of Ornstein-Uhlenbeck type $q = 1$ and $r = 1/\lambda$.
Then we obtain for $d = 1$ using (3.59) and (3.73)

$$\begin{aligned}
\mathbb{P}[M_\Omega \leq b] &= \mathbb{P}[m_\Omega \geq -b] \\
&= \mathbb{P}\left[\hat{M}_\Omega \leq \exp\{b\}\right] \\
&= \mathbb{P}[\hat{m}_\Omega \geq \exp\{-b\}] \\
&\approx \exp\{-\mu_b|\Omega|\} \\
&= \exp\left\{-f_0(b)\cdot\lambda^{-1}b|\Omega|\right\}.
\end{aligned} \tag{3.74}$$

and for $d = 2$ using (3.69) and (3.73)

$$\begin{aligned}
\exp\left\{-f_0(b)\cdot\lambda^{-2}b^3|\Omega|\right\} &\lesssim \mathbb{P}[M_\Omega \leq b] \\
&= \mathbb{P}[m_\Omega \geq -b] \\
&= \mathbb{P}\left[\hat{M}_\Omega \leq \exp\{b\}\right] \\
&= \mathbb{P}[\hat{m}_\Omega \geq \exp\{-b\}] \\
&\lesssim \exp\left\{-f_0(b)\cdot\lambda^{-2}b^3\cdot\frac{|\Omega|}{2}\right\}.
\end{aligned} \tag{3.75}$$

This finishes the proof.
Similar bounds (using inequality (3.69)) can also be given for $d = 3$.

□

This means that we have got an (as we will see by the numerical results in the next section) close approximation for the probability distribution of the maxima of fields

of Ornstein-Uhlenbeck type in the one dimensional case and close upper and lower bounds for it in higher dimensional cases. (To see how close in practice, please consider Table 3.3.) In all dimensions these bounds are significantly more accurate than using the bounds for general random fields (which do not take possible covariances into consideration) for our particular fields (as also shown by the numerical results in the following section). Similar bounds can also be found for Gaussian random fields with some other covariance functions.

3.5 Numerical Results

It is now interesting to compare the quality of these estimates with the numerical values, that we obtain for fields of Ornstein-Uhlenbeck type. This is done by finding the maxima of a large number of realisations of the field and computing the probability distribution of the maximum for this set of realisations. These numerically obtained values are then compared with the general expectation estimates from Section 3.3 (i.e. ignoring the covariance function), as well as with the improved results obtained using the Poisson clumping heuristics (see Section 3.4).

In the following table let $\hat{\mathbb{P}} := \mathbb{P}\left[\hat{M}_\Omega \leq \exp\{b\}\right]$, i.e. the probability that the maximum of a given lognormal field of given variance σ and correlation length λ is less or equal to a given value $\exp\{b\}$. The table now contains the values $\exp\{b\}$, such that $\hat{\mathbb{P}} = y\%$, where y is chosen to be 95, 99 and 99.9.
We get in the one-dimensional case using 100,000 realisations of a field on $\Omega = (0,1)$, with $N = 511$, $h = 1/N$ and $\lambda = 10h$ the following table, where "Estimated" means the value of $\exp(b)$ predicted by the theory in Section 3.4, whereas "Computed" means the result of the simulation experiments.

Var. σ^2	$\hat{\mathbb{P}} = 95\%$		$\hat{\mathbb{P}} = 99\%$		$\hat{\mathbb{P}} = 99.9\%$	
	Estimated	**Computed**	**Estimated**	**Computed**	**Estimated**	**Computed**
1	4.59×10^1	3.47×10^1	7.06×10^1	5.37×10^1	1.21×10^2	9.20×10^1
2	2.25×10^2	1.49×10^2	4.12×10^2	2.76×10^2	8.79×10^2	5.91×10^2
4	2.11×10^3	1.22×10^3	4.99×10^3	2.93×10^3	1.46×10^4	8.46×10^3
8	5.04×10^4	2.30×10^4	1.70×10^5	8.01×10^4	7.73×10^5	3.77×10^5
16	4.46×10^6	1.49×10^6	2.49×10^7	8.80×10^7	2.13×10^8	6.91×10^7
32	2.54×10^9	5.42×10^8	2.88×10^{10}	6.12×10^9	5.98×10^{11}	1.07×10^{11}

Table 3.1: Estimates based on Poisson clumping heuristics for $\Omega = (0,1)$, $N = 511$, $h = 1/N$ and $\lambda = 10h$ (based on (3.74))

This shows that the theoretically expected values for this case are pretty good and get only slightly worse for increasing variance σ^2. The main reason for these (small) errors lies in the fact, that the field for the numerical experiments was discretised, i.e. assumed to be piecewise constant on each fine grid element. By doing this, possible maxima are slightly averaged out. We can therefore expect that the bounds would be even better, if we considered the exact field in our numerical experiments as well. However the numerical experiments here show that the error by the discretisation seems to be small enough to be neglected.

Now let us consider the two dimensional case and let us first of all check the quality of the bounds that we derived for general random fields ("Theoretical bound", see Section 3.3). Consider therefore 10,000 realisations of fields of Ornstein-Uhlenbeck type with $\Omega = (0,1)^2$, $\sigma^2 = 4$ and $h = 1/32$ and various values for λ and compare the numerically obtained values for $\mathbb{E}\left[\max_{\tau \in \mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)\right]$ with the theoretical upper bounds for these values (see Table 3.2).

λ	**Num. obtained** $\mathbb{E}\left[\max_{\tau \subset \bar{\Omega}}(\alpha^\tau)\right]$	**Theor. bound for** $\mathbb{E}\left[\max_{\tau \subset \bar{\Omega}}(\alpha^\tau)\right]$
1/16	$2.39 \cdot 10^2$	$2.46 \cdot 10^3$
1/8	$1.33 \cdot 10^2$	$2.46 \cdot 10^3$
1/4	$6.31 \cdot 10^1$	$2.46 \cdot 10^3$
1/2	$4.50 \cdot 10^1$	$2.46 \cdot 10^3$

Table 3.2: General estimates for different correlation lengths in 2D for $\sigma^2 = 4$ and $h = 1/32$

We see that the bounds are acceptable for fields of short correlation length. However they are not very accurate, when smoother fields, i.e. fields with large correlation length with respect to the size of the domain, are chosen as these bounds do not take the covariance of the fields into consideration.

Next we would like to find out how good the (upper and lower) bounds, that we derived for fields of Ornstein-Uhlenbeck type in Section 3.4, perform in practice in 2D. Consider therefore 1,000 fields with $\Omega = (0,1)^2$, $h = 1/32$ and $\lambda = 1/32$. We then compare the numerically obtained values $\exp\left\{\tilde{b}\right\}$ such that $\mathbb{P}\left[\hat{M}_\Omega \leq \exp\left\{\tilde{b}\right\}\right] = \hat{\mathbb{P}}$, for $\hat{\mathbb{P}} = 0.9$, 0.925, 0.95, 0.975, 0.99 and 0.999, with the theoretically obtained lower bounds $\exp\{b\}$ and upper bounds $\exp\{B\}$ (based on inequality (3.75), see Table 3.3).

The upper and lower bounds here are not very far apart and therefore represent the true behaviour of the probability distribution of the maxima very well. The significance of the correlation length on this behaviour is very well represented in the bounds and we can use the bounds for the probability distributions of the condition numbers of the unpreconditioned and preconditioned stiffness matrices in the next few chapters to get good practical estimates for the convergence rates of our solvers.

λ	Prob. $\hat{\mathbb{P}}$	Theoretical lower bound $\exp\{\mathbf{b}\}$	Numerical value $\exp\left\{\tilde{\mathbf{b}}\right\}$	Theoretical upper bound $\exp\{\mathbf{B}\}$
1/16	90%	$4.46 \cdot 10^2$	$4.57 \cdot 10^2$	$5.80 \cdot 10^2$
	92.5%	$5.01 \cdot 10^2$	$5.33 \cdot 10^2$	$6.48 \cdot 10^2$
	95%	$5.86 \cdot 10^2$	$6.41 \cdot 10^2$	$7.51 \cdot 10^2$
	97.5%	$7.53 \cdot 10^2$	$8.73 \cdot 10^2$	$9.56 \cdot 10^2$
	99%	$1.03 \cdot 10^3$	$1.19 \cdot 10^3$	$1.29 \cdot 10^3$
	99.9%	$2.21 \cdot 10^3$	$2.51 \cdot 10^3$	$2.57 \cdot 10^3$
1/8	90%	$2.50 \cdot 10^2$	$2.59 \cdot 10^2$	$3.37 \cdot 10^2$
	92.5%	$2.85 \cdot 10^2$	$3.02 \cdot 10^2$	$3.82 \cdot 10^2$
	95%	$3.41 \cdot 10^2$	$3.68 \cdot 10^2$	$4.51 \cdot 10^2$
	97.5%	$4.53 \cdot 10^2$	$4.99 \cdot 10^2$	$5.89 \cdot 10^2$
	99%	$6.40 \cdot 10^2$	$7.89 \cdot 10^2$	$8.18 \cdot 10^2$
	99.9%	$1.39 \cdot 10^3$	$1.68 \cdot 10^3$	$1.72 \cdot 10^3$
1/4	90%	$1.25 \cdot 10^2$	$1.28 \cdot 10^2$	$1.80 \cdot 10^2$
	92.5%	$1.47 \cdot 10^2$	$1.62 \cdot 10^2$	$2.08 \cdot 10^2$
	95%	$1.82 \cdot 10^2$	$2.19 \cdot 10^2$	$2.53 \cdot 10^2$
	97.5%	$2.54 \cdot 10^2$	$3.02 \cdot 10^2$	$3.43 \cdot 10^2$
	99%	$3.77 \cdot 10^2$	$4.50 \cdot 10^2$	$4.95 \cdot 10^2$
	99.9%	$8.84 \cdot 10^2$	$1.09 \cdot 10^3$	$1.11 \cdot 10^3$
1/2	90%	$4.94 \cdot 10^1$	$5.13 \cdot 10^1$	$8.24 \cdot 10^1$
	92.5%	$6.27 \cdot 10^1$	$6.69 \cdot 10^1$	$9.96 \cdot 10^1$
	95%	$8.38 \cdot 10^1$	$9.05 \cdot 10^1$	$1.27 \cdot 10^2$
	97.5%	$1.28 \cdot 10^2$	$1.54 \cdot 10^2$	$1.83 \cdot 10^2$
	99%	$2.05 \cdot 10^2$	$2.53 \cdot 10^2$	$2.81 \cdot 10^2$
	99.9%	$5.40 \cdot 10^2$	$6.91 \cdot 10^2$	$6.96 \cdot 10^2$

Table 3.3: Theoretical and computed probability distributions for $\sigma^2 = 4$ and $h = 1/32$ (based on (3.75))

3.6 Summary

For general random fields as well as random fields with weak correlation, simple expectation bounds for their maxima and minima were found by norm inequalities. These bounds however proved to be over-pessimistic for fields with large correlation length. Although no general theory is available for such fields, recent ideas on Poisson clumping heuristics were used to get significantly better theoretical approximations for the probability distributions of the field maxima of one-dimensional fields and good upper and lower bounds for the probability distributions for higher dimensional fields.
The same is not only true for the maxima and minima themselves, but also for the maxima of local and global ratios of the field, as it will be shown and used in the following chapters. These bounds make it possible to give good predictions on the condition numbers of the stiffness matrices $A(\alpha)$ for groundwater flow problems (see Chapter 4), as well as of systems, that were preconditioned using domain decomposition (see Chapters 5 and 6).
Using these matrices in an iterative method like conjugate gradients, resp. preconditioned conjugate gradients, it is shown in Appendix B, that the number of iterations depend directly on these condition numbers and is proportional to $\frac{\sqrt{\kappa(A(\alpha))}-1}{\sqrt{\kappa(A(\alpha))}+1}$ in the unpreconditioned case and to $\frac{\sqrt{\kappa(M^{-1}A(\alpha))}-1}{\sqrt{\kappa(M^{-1}A(\alpha))}+1}$ if preconditioned with a matrix M^{-1}. The maxima estimates therefore also help to allow accurate predictions on the number of iterations the solver will need in average, as well as worst case analysis in 90%, 95%, etc. of the fields. Therefore the computational effort can be predicted in advance, which will allow users of the methods to develop a cost/risk analysis even before starting the computations.

Chapter 4

The unpreconditioned system

"Facts are meaningless. You could use facts to prove anything that's even remotely true!"

— Homer Simpson (Fictional character from the animated television series "The Simpsons" , created by Matt Groening) —

4.1 Convergence estimates for the unpreconditioned system

Let $A(\alpha)$ be the *stiffness matrix* (see (2.26)) with

$$\begin{aligned} A_{ij}(\alpha) &= \int_{\Omega} \alpha \nabla \phi_i . \nabla \phi_j \\ &= \sum_{\tau : \tau \subseteq \text{supp}(\phi_i) \cap \text{supp}(\phi_j)} \int_{\tau} \alpha \nabla \phi_i . \nabla \phi_j, \end{aligned} \tag{4.1}$$

where ϕ_i and ϕ_j are the basis functions (as defined in Subsection 2.2.2), chosen from the given subspace, i.e. in our case the subspace of piecewise linear functions, on a uniform triangulation of $\Omega = (0,1) \times (0,1)$ (see Figure 4-1).

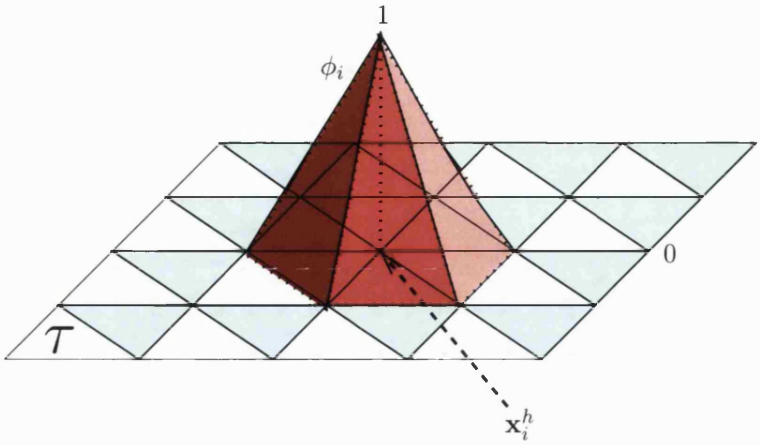


Figure 4-1: Grid with basis function ϕ_i

For each element $\tau \subset \bar{\Omega}$ define

$$\alpha^\tau := \alpha\left(\mathbf{m}^\tau\right), \tag{4.2}$$

where $\mathbf{m}^\tau$ is the centroid of τ.
(Realise that if the triangulation $\mathcal{T}^h(\bar{\Omega})$ is chosen fine enough with respect to the smoothness of the permeability field α, then $\alpha^\tau \approx \frac{1}{|\tau|}\int_\tau \alpha.$)
Then the matrix A is a function of the set of values $\boldsymbol{\alpha} := \{\alpha^\tau : \tau \in \mathcal{T}^h(\bar{\Omega})\}$, i.e. $A(\boldsymbol{\alpha})$, and given by the entries

$$A_{ij}(\boldsymbol{\alpha}) := \sum_{\tau:\tau\subseteq \mathrm{supp}(\phi_i)\cap\mathrm{supp}(\phi_j)} \alpha^\tau \int_\tau \boldsymbol{\nabla}\phi_i.\boldsymbol{\nabla}\phi_j, \tag{4.3}$$

i.e.

$$A_{ij}(\boldsymbol{\alpha}) = \sum_{\tau:\tau\subseteq \mathrm{supp}(\phi_i)\cap\mathrm{supp}(\phi_j)} \alpha^\tau A_{ij}^\tau(\mathbf{1}),$$

where $A^\tau(\mathbf{1})$ is the element stiffness matrix for element τ and coefficient vector $\mathbf{1}$.

On the basis of the simple results on condition numbers from Appendix A, we now would like to estimate the expected condition number of $A(\boldsymbol{\alpha})$ itself and of $A(\boldsymbol{\alpha})$ preconditioned by some matrix M^{-1}.

Let $\mathcal{N}^h(\Omega)$ denote the freedoms in Ω. It is now the easy to prove the following theorem.

THEOREM 4.1
Let $A(\mathbf{1})$ be the constant coefficient matrix and $A(\boldsymbol{\alpha})$ the matrix in (4.3). Then we

have

$$\lambda_{\max}\{A(\boldsymbol{\alpha})\} \leq \{\max_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)\}\lambda_{\max}\{A(\mathbf{1})\}, \tag{4.4}$$

$$\lambda_{\min}\{A(\boldsymbol{\alpha})\} \geq \{\min_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)\}\lambda_{\min}\{A(\mathbf{1})\}, \tag{4.5}$$

and

$$\kappa\left(A(\boldsymbol{\alpha})\right) \leq \frac{\max_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)}{\min_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)}\kappa\left(A(\mathbf{1})\right). \tag{4.6}$$

Proof:

Define the *local stiffness matrices* $A^\tau(\boldsymbol{\alpha})$ by

$$A^\tau_{ij}(\boldsymbol{\alpha}) := \int_\tau \alpha\boldsymbol{\nabla}\phi_i.\boldsymbol{\nabla}\phi_j = \left(\frac{1}{|\tau|}\int_\tau \alpha^\tau\right)\int_\tau \boldsymbol{\nabla}\phi_i.\boldsymbol{\nabla}\phi_j. \tag{4.7}$$

Let $\mathbf{x}^h_i \in \mathcal{N}^h(\Omega)$. Then we have for any $\mathbf{v}\in\mathbb{R}^n$,

$$\begin{aligned}
\mathbf{v}^T A(\boldsymbol{\alpha})\mathbf{v} &= \sum_i\sum_j v_i A_{ij}(\boldsymbol{\alpha})v_j \\
&= \sum_i\sum_j \sum_{\tau:\mathbf{x}^h_i,\mathbf{x}^h_j\in\tau} v_i\alpha^\tau A^\tau_{ij}(\mathbf{1})v_j \\
&= \sum_\tau \alpha^\tau \sum_{i:\mathbf{x}^h_i\in\tau}\sum_{j:\mathbf{x}^h_j\in\tau} v_i A^\tau_{ij}(\mathbf{1})v_j \\
&= \sum_\tau \alpha^\tau(\mathbf{v}^\tau)^T A^\tau(\mathbf{1})\mathbf{v}^\tau,
\end{aligned} \tag{4.8}$$

where $\mathbf{v}^\tau$ is the restriction of $\mathbf{v}$ to τ, i.e.

$$v^\tau_i = \begin{cases} v_i & \text{if } \mathbf{x}^h_i \in \tau, \\ 0 & \text{otherwise.} \end{cases}$$

The element stiffness matrices are positive semidefinite, that is:

$$(\mathbf{v}^\tau)^T A^\tau(\mathbf{1})\mathbf{v}^\tau \geq 0, \quad \forall \mathbf{v}\in\mathbb{R}^n.$$

Now using (4.8), we obtain

$$\begin{aligned}\mathbf{v}^T A(\boldsymbol{\alpha})\mathbf{v} &\geq \{\min_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)\}\sum_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\mathbf{v}^\tau)^T A^\tau(\mathbf{1})\mathbf{v}^\tau \\ &= \{\min_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)\}\mathbf{v}^T A(\mathbf{1})\mathbf{v}.\end{aligned} \tag{4.9}$$

As this is true for all vectors and therefore especially also for the unit eigenvector $\mathbf{v}$ corresponding to the minimum eigenvalue $\lambda_{\min}(A(\mathbf{1}))$

$$\mathbf{v}^T A(\boldsymbol{\alpha})\mathbf{v} \geq \{\min_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)\}\lambda_{\min}\{A(\mathbf{1})\}\mathbf{v}^T\mathbf{v}$$

and so

$$\lambda_{\min}\{A(\boldsymbol{\alpha})\} \geq \{\min_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)\}\lambda_{\min}\{A(\mathbf{1})\}. \tag{4.10}$$

Similarly, with the unit eigenvector corresponding to the maximum eigenvalue $\lambda_{\max}(A(\mathbf{1}))$

$$\lambda_{\max}\{A(\boldsymbol{\alpha})\} \leq \{\max_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)\}\lambda_{\max}\{A(\mathbf{1})\}. \tag{4.11}$$

Finally

$$\kappa\left(A(\boldsymbol{\alpha})\right) = \frac{\lambda_{\max}\{A(\boldsymbol{\alpha})\}}{\lambda_{\min}\{A(\boldsymbol{\alpha})\}} \leq \frac{\max\limits_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)}{\min\limits_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)}\kappa\left(A(\mathbf{1})\right). \tag{4.12}$$

□

4.2 Probability estimates

Let α now be a realisation of the lognormal Gaussian random field given by (3.12), i.e.

$$\alpha(\mathbf{x}) = \exp(Z(\mathbf{x})), \tag{4.13}$$

where $Z(\mathbf{x})$ is a realisation of a Gaussian random field with mean 0 and let the stiffness matrix $A(\boldsymbol{\alpha})$ be defined as in (4.3).
As shown in the previous section, the condition number of the stiffness matrix (4.3) depends linearly on the ratio $\{\max_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)\}/\{\min_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)\}$. As $\kappa\left(A(\mathbf{1})\right)$ is independent of any random variables, this means that upper bounds for all probability estimates of the condition numbers of the unpreconditioned stiffness matrices depend linearly on estimates for this (global) ratio.

4.2.1 General fields

As in the previous chapter, we can find upper bounds for the expectation of the maximum of the ratio $\{\max_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)\}/\{\min_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)\}$ with the help of simple norm inequalities. We can then prove the following theorem.

THEOREM 4.2
There exists a constant C independent of α and h, such that the expectation of the condition number of the stiffness matrix $A(\boldsymbol{\alpha})$ for any isotropic Gaussian random field can be bounded by

$$\mathbb{E}\left[\kappa\left(A\left(\boldsymbol{\alpha}\right)\right)\right] \leq \begin{cases} C\exp\left\{2\sqrt{2}\sigma\sqrt{\ln\left(\#\mathcal{T}_\Omega^h\right)}\right\}h^{-2}, & \text{when } \ln\left(\#\mathcal{T}_\Omega^h\right) \geq 2\sigma^2, \\ C\exp\left\{2\sigma^2\right\}\left(\#\mathcal{T}_\Omega^h\right)h^{-2}, & \text{otherwise.} \end{cases} \tag{4.14}$$

Proof:
As shown in Lemma 3.10 it is

$$\mathbb{E}\left[\frac{\max_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)}{\min_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)}\right] \leq \exp\{2p\sigma^2\}\{\#\mathcal{T}_\Omega^h\}^{1/p}, \tag{4.15}$$

for any $p \geq 1$.
Let $f(p) := \exp\{2p\sigma^2\}\{\#\mathcal{T}_\Omega^h\}^{1/p}$. Minimising this function, we find an infimum of the function for $p^* = 2^{-1/2}\sigma^{-1}\ln\left(\#\mathcal{T}_\Omega^h\right)$, if $\ln\left(\#\mathcal{T}_\Omega^h\right) \geq 2\sigma^2$. Therefore

$$\mathbb{E}\left[\frac{\max_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)}{\min_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)}\right] \leq \begin{cases} \exp\left\{2\sqrt{2}\sigma\sqrt{\ln\left(\#\mathcal{T}_\Omega^h\right)}\right\}, & \text{when } \ln\left(\#\mathcal{T}_\Omega^h\right) \geq 2\sigma^2, \\ \exp\left\{2\sigma^2\right\}\left(\#\mathcal{T}_\Omega^h\right) & \text{otherwise.} \end{cases} \tag{4.16}$$

We see that for $\ln\left(\#\mathcal{T}_\Omega^h\right) = 2\sigma^2$ the two bounds are equal.
As under the given regularity assumptions on the triangulation there exists a constant C independent of h and the permeability field, such that $\kappa\left(A(1)\right) \leq Ch^{-2}$ (see for example in Johnson [79], Section 7.7). Now from Theorem 4.1

$$\mathbb{E}\left[\kappa\left(A\left(\boldsymbol{\alpha}\right)\right)\right] \leq \mathbb{E}\left[\frac{\max_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)}{\min_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)}\right]\cdot\kappa\left(A\left(\mathbf{1}\right)\right) \leq \mathbb{E}\left[\frac{\max_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)}{\min_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)}\right]\cdot Ch^{-2} \tag{4.17}$$

and inequality (4.14) follow. □

4.2.2 Random fields of Ornstein-Uhlenbeck type

From now on let the underlying random fields be of Ornstein-Uhlenbeck type, i.e. with covariance as given in (3.13). We then can derive bounds for the probability distribution of $\kappa(A(\alpha))$.
As in Chapter 3 denote

$$f_0(b) := (2\pi)^{-d/2}\sigma^{-1}\exp\left\{-\frac{1}{2}db^2/\sigma^2\right\}, \qquad b \in \mathbb{R}. \tag{4.18}$$

Using Poisson clumping heuristics, we managed to show in Chapter 3 (under the then given assumptions) for one dimensional fields (i.e. $d = 1$)

$$\begin{aligned}\mathbb{P}\left[\sup_{\mathbf{x}\in\Omega}(\alpha(\mathbf{x})) \le \exp\{b\}\right] = \mathbb{P}\left[\inf_{\mathbf{x}\in\Omega}(\alpha(\mathbf{x})) \ge \exp\{-b\}\right] \\ \approx \exp\left\{-f_0(b)\cdot\lambda^{-1}b|\Omega|\right\}\end{aligned} \tag{4.19}$$

and for two dimensional fields (i.e. $d = 2$)

$$\begin{aligned}\exp\left\{-f_0(b)\cdot\lambda^{-2}b^3|\Omega|\right\} &\lesssim \mathbb{P}\left[\sup_{\mathbf{x}\in\Omega}(\alpha(\mathbf{x})) \le \exp\{b\}\right] \\ &= \mathbb{P}\left[\inf_{\mathbf{x}\in\Omega}(\alpha(\mathbf{x})) \ge \exp\{-b\}\right] \\ &\lesssim \exp\left\{-f_0(b)\cdot\lambda^{-2}b^3\frac{|\Omega|}{2}\right\}.\end{aligned} \tag{4.20}$$

Since $\alpha(\mathbf{x}) > 0$ for all $\mathbf{x} \in \Omega$, we now know that

$$\begin{aligned}\mathbb{P}\left[\max\left(\sup_{\mathbf{x}\in\Omega}(\alpha(\mathbf{x})^2), \inf_{\mathbf{x}\in\Omega}(\alpha(\mathbf{x})^{-2})\right) \le \exp\{b\}\right] &\le \mathbb{P}\left[\frac{\sup_{\mathbf{x}\in\Omega}(\alpha(\mathbf{x}))}{\inf_{\mathbf{y}\in\Omega}(\alpha(\mathbf{y}))} \le \exp\{b\}\right] \\ &\le \mathbb{P}\left[\min\left(\sup_{\mathbf{x}\in\Omega}(\alpha(\mathbf{x})^2), \inf_{\mathbf{x}\in\Omega}(\alpha(\mathbf{x})^{-2})\right) \le \exp\{b\}\right]\end{aligned} \tag{4.21}$$

and, since $\alpha(\mathbf{x}) = \exp\{Z(\mathbf{x})\}$, $\mathbf{x} \in \Omega$, is a lognormal random field, where $Z(\mathbf{x})$ is a Gaussian random variable with mean 0 and variance σ^2, that

$$\mathbb{P}\left[\sup_{\mathbf{x}\in\Omega}(\alpha(\mathbf{x})^2) \le \exp\{b\}\right] = \mathbb{P}\left[\inf_{\mathbf{x}\in\Omega}(\alpha(\mathbf{x})^{-2}) \le \exp\{b\}\right]. \tag{4.22}$$

Therefore

$$\mathbb{P}\left[\max\left(\sup_{\mathbf{x}\in\Omega}\left(\alpha(\mathbf{x})^2\right), \inf_{\mathbf{x}\in\Omega}\left(\alpha(\mathbf{x})^{-2}\right)\right) \leq \exp\{b\}\right] = \mathbb{P}\left[\sup_{\mathbf{x}\in\Omega}\left(\alpha(\mathbf{x})^2\right) \leq \exp\{b\}\right] \cdot \mathbb{P}\left[\inf_{\mathbf{x}\in\Omega}\left(\alpha(\mathbf{x})^{-2}\right) \leq \exp\{b\}\right]$$
$$= \mathbb{P}\left[\sup_{\mathbf{x}\in\Omega}\left(\alpha(\mathbf{x})^2\right) \leq \exp\{b\}\right]^2 \tag{4.23}$$

and

$$\mathbb{P}\left[\min\left(\sup_{\mathbf{x}\in\Omega}\left(\alpha(\mathbf{x})^2\right), \inf_{\mathbf{x}\in\Omega}\left(\alpha(\mathbf{x})^{-2}\right)\right) \leq \exp\{b\}\right] = \mathbb{P}\left[\sup_{\mathbf{x}\in\Omega}\left(\alpha(\mathbf{x})^2\right) \leq \exp\{b\}\right] + \mathbb{P}\left[\inf_{\mathbf{x}\in\Omega}\left(\alpha(\mathbf{x})^{-2}\right) \leq \exp\{b\}\right]$$
$$- \mathbb{P}\left[\sup_{\mathbf{x}\in\Omega}\left(\alpha(\mathbf{x})^2\right) \leq \exp\{b\}\right] \cdot \mathbb{P}\left[\inf_{\mathbf{x}\in\Omega}\left(\alpha(\mathbf{x})^{-2}\right) \leq \exp\{b\}\right]$$
$$= 2\mathbb{P}\left[\sup_{\mathbf{x}\in\Omega}\left(\alpha(\mathbf{x})^2\right) \leq \exp\{b\}\right] - \mathbb{P}\left[\sup_{\mathbf{x}\in\Omega}\left(\alpha(\mathbf{x})^2\right) \leq \exp\{b\}\right]^2. \tag{4.24}$$

Furthermore if $\alpha(\mathbf{x}) = \exp\{Z(\mathbf{x})\}$, $\mathbf{x} \in \Omega$, we obtain $\alpha(\mathbf{x})^2 = \exp\{2Z(\mathbf{x})\}$. We know that $2Z(\mathbf{x})$ therefore is a Gaussian random variable with mean 0 and variance $4\sigma^2$. Let now f_1 denote the function corresponding to f_0, when σ^2 is replaced by $4\sigma^2$, i.e.

$$f_1(b) := 2^{-1}(2\pi)^{-d/2}\sigma^{-1}\exp\left\{-\frac{1}{8}db^2/\sigma^2\right\}, \qquad b \in \mathbb{R}. \tag{4.25}$$

Using (4.19) we then obtain for $d = 1$

$$\mathbb{P}\left[\sup_{\mathbf{x}\in\Omega}\left(\alpha(\mathbf{x})\right)^2 \leq \exp\{b\}\right] = \mathbb{P}\left[\inf_{\mathbf{x}\in\Omega}\left(\alpha(\mathbf{x})^{-2}\right) \leq \exp\{b\}\right]$$
$$\approx \exp\left\{-f_1(b)\cdot\lambda^{-1}b|\Omega|\right\} \tag{4.26}$$

and using (4.20) we obtain for $d = 2$

$$\exp\left\{-f_1(b)\cdot\lambda^{-2}b^3|\Omega|\right\} \lesssim \mathbb{P}\left[\sup_{\mathbf{x}\in\Omega}\left(\alpha(\mathbf{x})^2\right) \leq \exp\{b\}\right]$$
$$= \mathbb{P}\left[\inf_{\mathbf{x}\in\Omega}\left(\alpha(\mathbf{x})^{-2}\right) \leq \exp\{b\}\right] \tag{4.27}$$
$$\lesssim \exp\left\{-f_1(b)\cdot\lambda^{-2}b^3\frac{|\Omega|}{2}\right\}.$$

Combining (4.21) with (4.23) and (4.24) we obtain using (4.26) for $d = 1$

$$\begin{aligned}
&\exp\left\{-2f_1(b)\cdot\lambda^{-1}b|\Omega|\right\}\\
&\lesssim \mathbb{P}\left[\frac{\sup_{\mathbf{x}\in\Omega}(\alpha(\mathbf{x}))}{\inf_{\mathbf{y}\in\Omega}(\alpha(\mathbf{y}))}\leq \exp\{b\}\right]\\
&\lesssim 2\exp\left\{-f_1(b)\cdot\lambda^{-1}b|\Omega|\right\}-\exp\left\{-2f_1(b)\cdot\lambda^{-1}b|\Omega|\right\}
\end{aligned} \tag{4.28}$$

and using (4.27) for $d = 2$

$$\begin{aligned}
&\exp\left\{-2f_1(b)\cdot\lambda^{-2}b^3|\Omega|\right\}\\
&\lesssim \mathbb{P}\left[\frac{\sup_{\mathbf{x}\in\Omega}(\alpha(\mathbf{x}))}{\inf_{\mathbf{y}\in\Omega}(\alpha(\mathbf{y}))}\leq \exp\{b\}\right]\\
&\lesssim 2\exp\left\{-f_1(b)\cdot\lambda^{-2}b^3\frac{|\Omega|}{2}\right\}-\exp\left\{-2f_1(b)\cdot\lambda^{-2}b^3|\Omega|\right\}.
\end{aligned} \tag{4.29}$$

Using Theorem 4.1 we obtain

$$\begin{aligned}
&\mathbb{P}\left[\kappa\left(A\left(\boldsymbol{\alpha}\right)\right)\leq \exp\{b\}\,\kappa\left(A\left(\mathbf{1}\right)\right)\right]\\
&\geq \mathbb{P}\left[\left(\kappa\left(A\left(\boldsymbol{\alpha}\right)\right)\leq \frac{\sup_{\mathbf{x}\in\Omega}(\alpha(\mathbf{x}))}{\inf_{\mathbf{y}\in\Omega}(\alpha(\mathbf{y}))}\kappa\left(A\left(\mathbf{1}\right)\right)\right)\cap\left(\frac{\sup_{\mathbf{x}\in\Omega}(\alpha(\mathbf{x}))}{\inf_{\mathbf{y}\in\Omega}(\alpha(\mathbf{y}))}\leq \exp\{b\}\right)\right]\\
&= \mathbb{P}\left[\kappa\left(A\left(\boldsymbol{\alpha}\right)\right)\leq \frac{\sup_{\mathbf{x}\in\Omega}(\alpha(\mathbf{x}))}{\inf_{\mathbf{y}\in\Omega}(\alpha(\mathbf{y}))}\kappa\left(A\left(\mathbf{1}\right)\right)\right]\cdot\mathbb{P}\left[\frac{\sup_{\mathbf{x}\in\Omega}(\alpha(\mathbf{x}))}{\inf_{\mathbf{y}\in\Omega}(\alpha(\mathbf{y}))}\leq \exp\{b\}\right]\\
&= \mathbb{P}\left[\frac{\sup_{\mathbf{x}\in\Omega}(\alpha(\mathbf{x}))}{\inf_{\mathbf{y}\in\Omega}(\alpha(\mathbf{y}))}\leq \exp\{b\}\right]
\end{aligned} \tag{4.30}$$

Therefore we find the following bounds for the probability distribution of the condition number of the stiffness matrix $A(\alpha)$.

COROLLARY 4.3
For fields α of Ornstein-Uhlenbeck type (as defined in Chapter 3), we get for one dimensional problems

$$\exp\left\{-2f_1(b)\cdot\lambda^{-1}b|\Omega|\right\}\lesssim \mathbb{P}\left[\kappa\left(A\left(\boldsymbol{\alpha}\right)\right)\leq \exp\{b\}\,\kappa\left(A\left(\mathbf{1}\right)\right)\right] \tag{4.31}$$

and for two dimensional problems

$$\exp\left\{-2f_1(b)\cdot\lambda^{-2}b^3|\Omega|\right\}\lesssim \mathbb{P}\left[\kappa\left(A\left(\boldsymbol{\alpha}\right)\right)\leq \exp\{b\}\,\kappa\left(A\left(\mathbf{1}\right)\right)\right] \tag{4.32}$$

4.3 Numerical results

As shown in Section 4.1, the condition numbers of the unpreconditioned system can in the worst case grow linearly with $\{\max_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)\}/\{\min_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)\}$. We therefore first of all would like to test how well the theoretical bounds for the expectation $\mathbb{E}\left[\max_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)\,/\,\min_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)\right]$ for general fields (based on simple norm inequalities), resp. for $\hat{\mathbb{P}}(b) := \mathbb{P}\left[\{\max_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)\}/\{\min_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)\}\leq\exp\{b\}\right]$ (based on Poisson clumping heuristics) compare with the probability distributions obtained in practice when considering a large number of random field realisations.

We computed 1,000 realisations of fields with $h = 1/32$ with covariance of Ornstein-Uhlenbeck type with different correlation lengths λ and obtained Table 4.1 for the general field estimates.

λ	**Numerically obtained**	**Theoretical bound (4.16) for**
	$\mathbb{E}\left[\max_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)\,/\,\min_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)\right]$	$\mathbb{E}\left[\max_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)\,/\,\min_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)\right]$
1/32	$4.23\cdot10^5$	$6.25\cdot10^9$
1/16	$1.41\cdot10^5$	$6.25\cdot10^9$
1/8	$2.35\cdot10^4$	$6.25\cdot10^9$
1/4	$3.48\cdot10^3$	$6.25\cdot10^9$
1/2	$4.61\cdot10^2$	$6.25\cdot10^9$

Table 4.1: General estimates for different correlation lengths in 2D for $\sigma^2 = 4$ and $h = 1/32$

Similar to the results for expectations of field maxima of lognormal random fields without taking their covariance into consideration (Table 3.2 in Chapter 3), these theoretical bounds are extremely pessimistic (even for short correlation length) and are therefore not very useful for estimates in practice. In the previous chapter we have however also developed bounds for probability distributions for maxima of fields with covariance of Ornstein Uhlenbeck type (see (4.29)). We therefore compare the theoretical lower bounds, $\exp\{b\}$, and upper bounds, $\exp\{B\}$, from this inequality with the numerically obtained values, $\exp\left\{\widetilde{b}\right\}$, from 1000 fields for probabilities of 85% and more.

The upper and lower bounds are very close to the numerically obtained distributions for short as well as for long correlation length for probabilities $\hat{\mathbb{P}}(b)\geq 0.85$. Again we therefore see that the bounds obtained using Poisson clumping heuristics are very useful for estimates in practice.

λ	Prob. $\hat{\mathbb{P}}[\mathrm{b}]$	Theoretical lower bound $\exp\{\mathrm{b}\}$	Numerical value $\exp\left\{\tilde{\mathrm{b}}\right\}$	Theoretical upper bound $\exp\{\mathrm{B}\}$
1/16	85%	$1.17 \cdot 10^5$	$1.85 \cdot 10^5$	$1.06 \cdot 10^6$
	87.5%	$1.22 \cdot 10^5$	$2.26 \cdot 10^5$	$1.20 \cdot 10^6$
	90%	$1.26 \cdot 10^5$	$2.75 \cdot 10^5$	$1.40 \cdot 10^6$
	92.5%	$1.31 \cdot 10^5$	$3.58 \cdot 10^5$	$1.70 \cdot 10^6$
	95%	$1.36 \cdot 10^5$	$5.38 \cdot 10^5$	$2.21 \cdot 10^6$
	97.5%	$1.42 \cdot 10^5$	$8.97 \cdot 10^5$	$3.39 \cdot 10^6$
	99%	$1.48 \cdot 10^5$	$1.78 \cdot 10^6$	$5.80 \cdot 10^6$
1/8	85%	$3.36 \cdot 10^4$	$3.67 \cdot 10^4$	$4.07 \cdot 10^5$
	87.5%	$3.51 \cdot 10^4$	$3.99 \cdot 10^4$	$4.69 \cdot 10^5$
	90%	$3.66 \cdot 10^4$	$4.40 \cdot 10^4$	$5.54 \cdot 10^5$
	92.5%	$3.83 \cdot 10^4$	$5.80 \cdot 10^4$	$6.83 \cdot 10^5$
	95%	$4.01 \cdot 10^4$	$8.26 \cdot 10^4$	$9.07 \cdot 10^5$
	97.5%	$4.21 \cdot 10^4$	$1.43 \cdot 10^5$	$1.44 \cdot 10^6$
	99%	$4.34 \cdot 10^4$	$2.86 \cdot 10^5$	$2.56 \cdot 10^6$
1/4	85%	$7.14 \cdot 10^3$	$7.41 \cdot 10^3$	$1.41 \cdot 10^5$
	87.5%	$7.54 \cdot 10^3$	$7.99 \cdot 10^3$	$1.65 \cdot 10^5$
	90%	$7.98 \cdot 10^3$	$8.93 \cdot 10^3$	$1.99 \cdot 10^5$
	92.5%	$8.46 \cdot 10^3$	$9.84 \cdot 10^3$	$2.51 \cdot 10^5$
	95%	$8.98 \cdot 10^3$	$1.09 \cdot 10^4$	$3.43 \cdot 10^5$
	97.5%	$9.56 \cdot 10^3$	$1.78 \cdot 10^4$	$5.70 \cdot 10^5$
	99%	$9.93 \cdot 10^3$	$3.97 \cdot 10^4$	$1.07 \cdot 10^6$
1/2	85%	$5.92 \cdot 10^2$	$6.64 \cdot 10^2$	$4.17 \cdot 10^4$
	87.5%	$6.71 \cdot 10^2$	$7.83 \cdot 10^2$	$5.01 \cdot 10^4$
	90%	$7.64 \cdot 10^2$	$9.10 \cdot 10^2$	$6.23 \cdot 10^4$
	92.5%	$8.44 \cdot 10^2$	$1.11 \cdot 10^3$	$8.13 \cdot 10^4$
	95%	$9.32 \cdot 10^2$	$1.50 \cdot 10^3$	$1.16 \cdot 10^5$
	97.5%	$1.24 \cdot 10^3$	$2.20 \cdot 10^3$	$2.05 \cdot 10^5$
	99%	$1.73 \cdot 10^3$	$5.70 \cdot 10^3$	$4.10 \cdot 10^5$

Table 4.2: Theoretical and computed probability distributions for $\sigma^2 = 4$ and $h = 1/32$

4.4 Summary

The main aim of this short chapter was to show how ill-conditioned the matrix $A(\boldsymbol{\alpha})$ with entries as given in (4.1) can be for strongly heterogeneous coefficients. We therefore wrote the global stiffness matrix $A(\boldsymbol{\alpha})$ of the problem as the sum of local stiffness matrices. Based on this extension we managed to prove that the condition number can in the worst case depend linearly on the global ratio $\{\max_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)\}/\{\min_{\tau\in\mathcal{T}^h(\bar{\Omega})}(\alpha^\tau)\}$, which as mentioned in the introduction of the groundwater problem can be of the order of 10^9 or more for real world problems (see Section 2.5).
Furthermore we managed to give estimates on the expectation of this ratio for general random fields and estimates for its probability distribution for fields with covariance of Ornstein-Uhlenbeck type. Especially the latter can be very precise (see therefore also the numerical results in Section 3.5) and can be used in practice to give estimates on computation times for algorithms involving matrix $A(\boldsymbol{\alpha})$.
As we want to solve the discretised version of problem (2.13,2.14) iteratively and as we can show (see Appendix B) that the convergence of the unpreconditioned conjugate gradient method is given by the ratio

$$\left(\sqrt{\kappa\left(A(\boldsymbol{\alpha})\right)}-1\right)/\left(\sqrt{\kappa\left(A(\boldsymbol{\alpha})\right)}+1\right),$$

we will use the next three chapters of this thesis to find a preconditioner M^{-1}, which reduces the dependence of $M^{-1}A(\boldsymbol{\alpha})$ on h and especially on α. Using a preconditioned version of the conjugate gradient method will then lead to a convergence rate

$$\left(\sqrt{\kappa\left(M^{-1}A(\boldsymbol{\alpha})\right)}-1\right)/\left(\sqrt{\kappa\left(M^{-1}A(\boldsymbol{\alpha})\right)}+1\right),$$

that will be close to 0, if $\kappa\left(M^{-1}A(\boldsymbol{\alpha})\right)$ is close to 1.
Since for the groundwater problem considered $\kappa\left(M^{-1}A(\boldsymbol{\alpha})\right)$ is often very large, it is very essential to find good preconditioners. Attempts to find such preconditioners will be described in the following chapters.

Chapter 5

Domain decomposition with linear interpolation

"*Nothing is particularly hard if you divide it into small jobs.*"

— Henry Ford (American industrialist and pioneer of the assembly line production method, 1863-1947) —

"*There are two kinds of people in this world - those who divide everything into two and those who don't.*"

— Robert Benchley (American actor, author and humorist, 1889-1945) —

5.1 Abstract theory of Schwarz methods

5.1.1 Introduction

As it was shown in the previous chapter, it is very essential to find a "good" preconditioner for the finite element discretisation of the problem

$$-\mathrm{div}\left(\alpha\left(\mathbf{x}\right)\nabla u\left(\mathbf{x}\right)\right)=f\left(\mathbf{x}\right)\ \text{ for } \mathbf{x}\in\Omega, \tag{5.1}$$

with

$$u\left(\mathbf{x}\right)=0 \text{ for } \mathbf{x}\in\Gamma, \tag{5.2}$$

especially, if α is varying strongly on the domain.

But what exactly would we like a "good" preconditioner to be like?

(i) **Reduction of dependence on α:** As shown before, the condition number of the unpreconditioned system depends in the worst case linearly on the global ratio $\max_{\mathbf{x},\mathbf{y}\in\Omega}(\alpha(\mathbf{x})/\alpha(\mathbf{y}))$. For "smooth" random fields, i.e. fields with a large correlation length, a reduction of the condition number would be achieved, if we can get this dependence down to a more local ratio of values of α.

(ii) **Potential of parallelization:** We would ideally like to be able to solve problems on very fine grids to best resolve the underlying permeability field. Furthermore we have to solve the problem for a large number of fields to be then able to apply Monte Carlo methods afterwards. This can be computationally very expensive, might need a lot of computer memory and is often not possible on a single processor. We therefore would prefer to use a method that can be parallelised easily without too much communication between the processors. So the different processes should be independent and the amount of data transfer should be minimised.

(iii) **Reliability:** Our preconditioner should have well-understood convergence properties. For problems with realisations of random fields as permeability fields this might mean knowledge of how many iterations we have to expect in average, or an upper bound in 90%, 95% or 99% of the cases.

(iv) **Flexibility:** Complex and irregular geometries should be treated locally (i.e. if the method is parallelised on one processor without need of communication), so that the preconditioner can also handle more complicated domains than the unit square, resp. unit cube of our test problems. Ideally the preconditioner should handle many different domains and grids.

To understand the preconditioners discussed later in this chapter, let us first of all look at a method developed by Hermann Amandus Schwarz in 1870 (see Schwarz [118]), the so called Schwarz alternating method. This method is thought to be the first attempt to solve a problem on a complicated domain, by solving local subproblems.
Consider therefore a decomposition of the domain Ω into two overlapping subdomains $\widetilde{\Omega_1}$ and $\widetilde{\Omega_2}$, which are supposed to cover Ω completely. Now let $\mathcal{A}$ be the operator such that

$$\mathcal{A} : u \longmapsto -\operatorname{div}(\alpha \nabla u(\mathbf{x})), \quad \forall u \in C^2(\Omega).$$

Problem (5.1,5.2) then can be written as

$$\begin{aligned} \mathcal{A}u &= f \text{ in } \Omega, \\ u &= 0 \text{ on } \Gamma, \text{ where } \Gamma \text{ is the boundary of } \Omega. \end{aligned} \tag{5.3}$$

Now let $\Gamma_i := \partial\widetilde{\Omega_i} \backslash \partial\Omega, \quad i = 1, 2.$
An example for a domain decomposition of this kind is given in Figure 5-1.

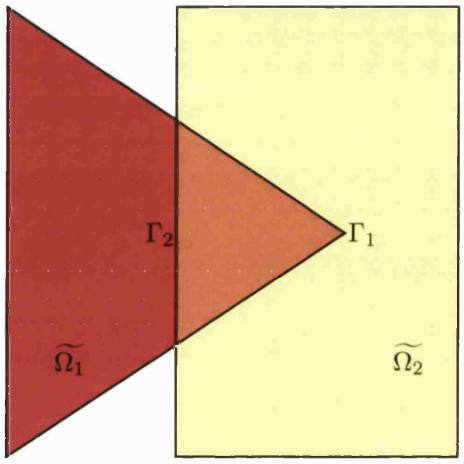


Figure 5-1: Two overlapping subdomains

Let $\mathcal{A}_1$ be the operator such that

$$\mathcal{A}_1 : u_1 \longmapsto -\text{div}\left(\alpha \boldsymbol{\nabla} u_1(\mathbf{x})\right), \quad \forall u_1 \in C^2\left(\widetilde{\Omega_1}\right),$$

and let $\mathcal{A}_2$ be the operator such that

$$\mathcal{A}_2 : u_2 \longmapsto -\text{div}\left(\alpha \boldsymbol{\nabla} u_2(\mathbf{x})\right), \quad \forall u_2 \in C^2\left(\widetilde{\Omega_2}\right).$$

With this notation we get

Schwarz alternating method:

Choose an initial guess u_2^0.
For $k = 1, 2, \ldots$ *solve*

$$\begin{cases} \mathcal{A}_1 u_1^k = f & \text{in } \widetilde{\Omega_1}, \\ u_1^k = 0 & \text{on } \partial\widetilde{\Omega_1}\backslash\Gamma_1, \\ u_1^k = u_2^{k-1}|_{\Gamma_1} & \text{on } \Gamma_1, \end{cases} \tag{5.4}$$

and

$$\begin{cases} \mathcal{A}_2 u_2^k = f & \text{in } \widetilde{\Omega_2}, \\ u_2^k = 0 & \text{on } \partial\widetilde{\Omega_2}\backslash\Gamma_2, \\ u_2^k = u_1^{k-1}|_{\Gamma_2} & \text{on } \Gamma_2. \end{cases} \tag{5.5}$$

This means we solve an elliptic boundary problem on each subdomain $\widetilde{\Omega_j}$, $j = 1, 2$, with u given exactly on Γ and updated with values on the artificial boundary Γ_j. The approximate solution at the k-th iteration step is then defined by

$$u^k(\mathbf{x}) = \begin{cases} u_2^k(\mathbf{x}) & \text{if } \mathbf{x} \in \widetilde{\Omega_2}, \\ u_1^k(\mathbf{x}) & \text{if } \mathbf{x} \in \Omega\backslash\widetilde{\Omega_2}. \end{cases} \tag{5.6}$$

Furthermore it can be shown that the iterates $\{u^k\}$ converge to the true solution u (see Lions [91], Section I.2).
This simple description might give a first idea of a method of splitting a large problem into two subproblems. The following subsection will show, how we can discretise this method and how we can use it as a preconditioner.

5.1.2 Additive Schwarz methods

In a similar way as done by Toselli and Widlund [128], Chapter 1, let us now introduce a *coarse grid* with elements of maximum diameter H, which defines a triangulation $\mathcal{T}^H(\bar{\Omega})$ with open *triangular subdomains* (in 2D), resp. *tetrahedral subdomains* (in 3D), $\Omega_1, \ldots, \Omega_p$. We then extend these subdomains to open *overlapping subdomains*, $\widetilde{\Omega_1}, \ldots, \widetilde{\Omega_p}$, with a minimum overlap of size βH, where $0 < \beta < 1$, i.e. any point on the boundary of an overlapping subdomain $\widetilde{\Omega_i}$, $i = 1, ..., p$, that is not also on the

boundary $\partial\Omega$ of Ω is at least βH away from its nearest point on the boundary of the non-overlapping subdomain Ω_i.

Assume that the fine grid is a subgrid of the coarse grid and that it triangulates the overlapping subdomains, $\widetilde{\Omega_j}$, $j \in \{1,\ldots,p\}$.

Then we define (similarly as in Section 2.2) on Ω

$$\mathcal{V}^h := \mathcal{S}_0^h(\Omega),$$

where $\mathcal{S}_0^h(\Omega) := \mathcal{S}^h(\Omega) \cap H_0^1(\Omega)$ and where $\mathcal{S}^h(\Omega)$ denotes the space of continuous piecewise linear functions with respect to $\mathcal{T}^h(\bar{\Omega})$ on Ω.

Define on each $\widetilde{\Omega_j}$

$$\mathcal{V}_j := \left\{ v_h \in \mathcal{V}^h : \ \operatorname{supp}(v_h) \subset \overline{\widetilde{\Omega_j}} \right\}, \tag{5.7}$$

where $\overline{\widetilde{\Omega_j}}$ denotes the closure of $\widetilde{\Omega_j}$.
Let $\mathcal{N}^h(\Omega)$ denote the interior nodes in Ω w.r.t. the fine grid $\mathcal{T}^h(\bar{\Omega})$ and let n_h be the cardinality of $\mathcal{N}^h(\Omega)$. In a similar way let $\mathcal{N}^H(\Omega)$ be the interior nodes in Ω w.r.t. the coarse grid $\mathcal{T}^H(\bar{\Omega})$ and let n_H be their cardinality. Finally let $\mathcal{N}^h(\widetilde{\Omega_j})$, $j \in \{1,\ldots,p\}$, be the set of interior fine grid nodes in $\widetilde{\Omega_j}$ and let n_j be their cardinality.

We shall now define the so called *extension operators* $\mathcal{R}_j^T$ as follows

$$\mathcal{R}_j^T : \mathcal{V}_j \to \mathcal{V}^h, \quad j \in \{1,\ldots,p\}.$$

If we represent a function $v_j \in \mathcal{V}_j$ by the vector $\mathbf{v_j}$ of its values at the freedoms in the interior of $\widetilde{\Omega_j}$, a map R_j^T (corresponding to the extension operator $\mathcal{R}_j^T$) is given by

$$\left(R_j^T \mathbf{v}_j\right)_l := \begin{cases} (\mathbf{v}_j)_l & \text{for } \mathbf{x}_l^h \in \mathcal{N}^h(\widetilde{\Omega_j}), \\ 0 & \text{for } \mathbf{x}_l^h \in \mathcal{N}^h(\Omega) \backslash \mathcal{N}^h(\widetilde{\Omega_j}). \end{cases} \tag{5.8}$$

Writing R_j^T in matrix form results in a $n_h \times n_j$ matrix. Its transpose R_j is often called a *local restriction matrix* and together with the global stiffness matrix $A := A(\boldsymbol{\alpha})$ they define

$$A_j := R_j A R_j^T, \tag{5.9}$$

a so called *local stiffness matrix*. A_j is therefore just a minor of A containing the values of the rows and columns of A that correspond to the nodes $\mathcal{N}^h(\widetilde{\Omega_j})$.

We will now come back to the two subdomain problem discussed in the previous section. Let $\mathbf{u}$ and $\mathbf{f}$ be the vectors of coefficients of u, resp. f written in the finite element basis and let $\mathbf{u}^k$ be the approximation of $\mathbf{u}$ at the k-th iteration step of the following algorithms and let $\mathbf{u}^{k+1/2}$ be a help vector, used as intermediate approximation of $\mathbf{u}$ in these algorithms. Furthermore split $\mathbf{f}$ into the parts $\mathbf{f}_1$, resp. $\mathbf{f}_2$ corresponding to the overlapping subdomains $\widetilde{\Omega_1}$, resp. $\widetilde{\Omega_2}$. We can then formulate a **discrete version of the Schwarz alternating method:**

Choose an initial guess $\mathbf{u}^0$.
For $k = 1, 2, \ldots$ *compute*

$$\mathbf{u}^{k+1/2} = \mathbf{u}^k + R_1^T A_1^{-1} R_1 \left(\mathbf{f} - A\mathbf{u}^k\right), \tag{5.10}$$

$$\mathbf{u}^{k+1} = \mathbf{u}^{k+1/2} + R_2^T A_2^{-1} R_2 \left(\mathbf{f} - A\mathbf{u}^{k+1/2}\right). \tag{5.11}$$

For large enough overlap, it can be shown that this algorithm converges with a rate independent of the mesh size h.

Based on this one can define an analogous **block Jacobi version**:

Choose an initial guess $\mathbf{u}^0$.
For $k = 1, 2, \ldots$ *compute*

$$\mathbf{u}^{k+1/2} = \mathbf{u}^k + R_1^T A_1^{-1} R_1 \left(\mathbf{f} - A\mathbf{u}^k\right), \tag{5.12}$$

$$\mathbf{u}^{k+1} = \mathbf{u}^{k+1/2} + R_2^T A_2^{-1} R_2 \left(\mathbf{f} - A\mathbf{u}^k\right). \tag{5.13}$$

The method is better parallelisable than the Schwarz alternating method, as the two subdomain solves can be performed at the same time and independently. However it is not guaranteed to converge anymore. If we now eliminate $\mathbf{u}^{j+1/2}$ this method simplifies to

Choose an initial guess $\mathbf{u}^0$.
For $k = 1, 2, \ldots$ *compute*

$$\mathbf{u}^{k+1} = \mathbf{u}^k + \left(R_1^T A_1^{-1} R_1 + R_2^T A_2^{-1} R_2\right)\left(\mathbf{f} - A\mathbf{u}^k\right), \tag{5.14}$$

which is a simple *Richardson iteration* on $A\mathbf{u} = \mathbf{f}$ with preconditioner

$$\hat{M}^{-1} := R_1^T A_1^{-1} R_1 + R_2^T A_2^{-1} R_2.$$

This is known as *one-level additive Schwarz preconditioner* for two overlapping subdomains. As $\hat{M}^{-1}A$ is symmetric with respect to the inner product induced by A, we can use this preconditioner inside a conjugate gradient method.

For p subdomains $\Omega_1, ..., \Omega_p$ this one level preconditioner can then be generalized to obtain

$$M_1^{-1} := \sum_{j=1}^{p} R_j^T A_j^{-1} R_j.$$

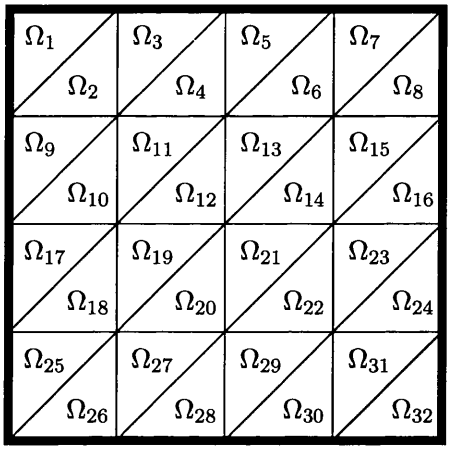


Figure 5-2: Domain with 32 nonoverlapping subdomains

Define $\bar{\Omega}_j$ to be the closure of Ω_j, $j = 1, ..., p$. For each coarse grid node $\mathbf{x}_i^H \in \mathcal{N}^H\left(\bar{\Omega}\right)$

and for each coarse grid element $\bar{\Omega}_j \in \mathcal{T}^H(\bar{\Omega})$ define the open subsets of Ω

$$\omega_i := \text{interior}\left(\bigcup_{j:\mathbf{x}_i^H \in \bar{\Omega}_j} \bar{\Omega}_j\right), \tag{5.15}$$

and

$$\omega_{\Omega_j} := \text{interior}\left(\bigcup_{k:\bar{\Omega}_k \in \mathcal{T}^H(\bar{\Omega}), \bar{\Omega}_k \cap \bar{\Omega}_j \neq \emptyset} \bar{\Omega}_k\right). \tag{5.16}$$

In this thesis we will consider (coarse grid) functions whose values will be determined by data at the nodes of the coarse grid. Thus the coarse grid space, $\mathcal{V}_0$, is a generalisation of the usual space of continuous piecewise linear functions on $\mathcal{T}^H(\bar{\Omega})$. More precisely we need the basis functions ψ_i^0, $i \in \{1, \ldots n_H\}$, of $\mathcal{V}_0$ to satisfy the following four conditions:

(i) $\psi_i^0 \in \mathcal{S}^h(\Omega)$, $\quad \psi_i^0\left(\mathbf{x}_j^H\right) = \delta_{i,j}$, for all $\mathbf{x}_j^H \in \mathcal{N}^H(\bar{\Omega})$,

(ii) $\text{supp}\left(\psi_i^0\right) \subset \bar{\omega}_i$,

(iii) $\sum_{\mathbf{x}_i^H \in \mathcal{N}^H(\bar{\Omega})} \psi_i^0(\mathbf{x}) = 1$

(iv) $\left\|\psi_i^0\right\|_{L_\infty(\Omega)} \leq 1$.

Due to condition (i) these basis functions are linearly independent. Based on these functions we define the space

$$\mathcal{V}_0 := \text{span}\left\{\psi_i^0 : \mathbf{x}_i^H \in \mathcal{N}^H(\Omega)\right\}. \tag{5.17}$$

By (i) and (ii) $\mathcal{V}_0$ is the span of all ψ_i^0 that vanish on the boundary $\partial\Omega$. It is therefore a subspace of $\mathcal{V}^h$.

Furthermore let $\boldsymbol{\psi}_1^0, \ldots, \boldsymbol{\psi}_{n_H}^0$ be the vectors of values of the coarse grid basis functions at the fine grid nodes.

We then define R_0^T to be the $n_h \times n_H$ matrix

$$R_0^T := \left[\boldsymbol{\psi}_1^0, \ldots, \boldsymbol{\psi}_{n_H}^0\right]. \tag{5.18}$$

As R_0^T maps function values at the coarse grid nodes to values at the fine grid nodes, it is often called a *global extension operator*, its transpose R_0 a *global restriction operator.*

Together with the global stiffness matrix A, they define

$$A_0 := R_0 A R_0^T, \tag{5.19}$$

the so called *coarse grid stiffness matrix.* Since the meshes $\mathcal{T}^h(\bar{\Omega})$ and $\mathcal{T}^H(\bar{\Omega})$ are nested and due to the nature of linear interpolation, A_0 is the stiffness matrix arising from discretising problem (5.1, 5.2) on the coarse mesh $\mathcal{T}^H(\bar{\Omega})$.

Using these notations, we define similarly to the *one level additive Schwarz preconditioner*

$$M_1^{-1} := \sum_{j=1}^{p} R_j^T A_j^{-1} R_j$$

the *two level additive Schwarz preconditioner*

$$M_{2,AS}^{-1} := \sum_{j=0}^{p} R_j^T A_j^{-1} R_j = M_1^{-1} + R_0^T A_0^{-1} R_0.$$

5.1.3 Basic properties of additive Schwarz preconditioners

With matrices $P_j := R_j^T A_j^{-1} R_j A$, $j = 0, ..., p$, the two level additive Schwarz preconditioner $M_{2,AS}^{-1}$ for (non-overlapping) subregions Ω_j (and overlapping subregions $\widetilde{\Omega_j}$) with $j \in \{1, \ldots, p\}$ is defined by

$$M_{2,AS}^{-1} A = \sum_{j=0}^{p} R_j^T A_j^{-1} R_j A = \sum_{j=0}^{p} P_j. \tag{5.20}$$

DEFINITION 5.1
For any function $v_h \in \mathcal{V}^h$, let $\mathbf{v}_h$ in this section and the next denote the vector of values of v_h at the fine grid nodes $\mathbf{x}_i^h \in \mathcal{N}^h(\Omega)$. (The vector $\mathbf{v}_h$ determines the function v_h uniquely, since v_h is piecewise linear with respect to the fine grid $\mathcal{T}^h(\bar{\Omega})$.)

We know that A is symmetric positive definite. Therefore as

$$\begin{aligned} P_j P_j &= R_j^T A_j^{-1} R_j A R_j^T A_j^{-1} R_j A \\ &= R_j^T A_j^{-1} A_j A_j^{-1} R_j A \\ &= R_j^T A_j^{-1} R_j A = P_j, \end{aligned} \tag{5.21}$$

the P_j, $j \in \{1, ..., p\}$, are projections. If we introduce the inner product $a(\mathbf{v}, \mathbf{w}) := \mathbf{w}^T A \mathbf{v}$, we obtain for any $v, w \in \mathcal{V}^h$

$$a\left(P_j \mathbf{v}, \mathbf{w}\right) = \mathbf{w}^T A P_j \mathbf{v} = \mathbf{w}^T A \left(R_j^T A_j^{-1} R_j A\right) \mathbf{v} = \left(R_j^T A_j^{-1} R_j A \mathbf{w}\right)^T A \mathbf{v} = a\left(\mathbf{v}, P_j \mathbf{w}\right). \tag{5.22}$$

Therefore the P_j are self adjoint with respect to the scalar product induced by $a(.,.)$. Furthermore P_j, $j = 1, ..., p$, projects V^h into the subspace V_j (as defined in (5.7)).

This means that for $u \in \mathcal{V}^h$ we have

$$\mathbf{w} := P_j \mathbf{u} \text{ (i.e. } w \in \mathcal{V}_j) \text{ satisfies } a\left(P_j \mathbf{u}, \mathbf{v}\right) = a(\mathbf{u}, \mathbf{v}) \quad \forall \mathbf{v} \text{ with } v \in \mathcal{V}_j.$$

To find bounds of the maximum and minimum eigenvalues of $M_1^{-1} A$, resp. $M_{2,AS}^{-1} A$, we can equivalently try to find bounds for the spectra of the sum of these orthogonal projections, i.e. of $\sum_{j=1}^{p} P_j$, resp. of $\sum_{j=0}^{p} P_j$.

Now consider orthogonal spaces in the $a(.,.)$ inner product among the (local) subspaces $\mathcal{V}_1, \ldots, \mathcal{V}_p$. To do so give each subspace a colour and let N_c be the *minimum number of colours* such that subspaces corresponding to neighboured subdomains (i.e. subdomains sharing at least a common point with a given subdomain) have different colours (see Figure 5-3).

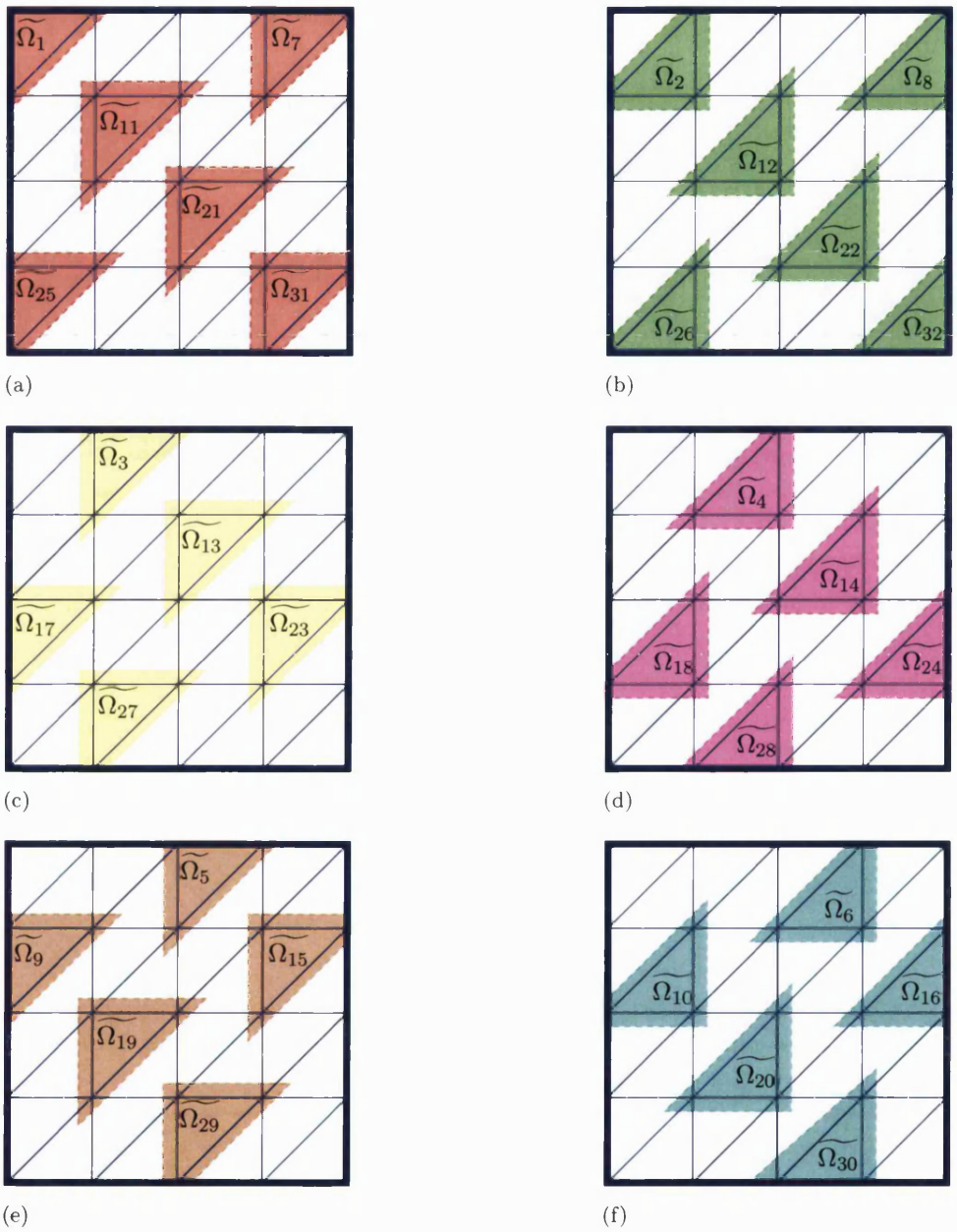


Figure 5-3: Colouring of the subdomains for a uniform coarse grid of 32 triangular subdomains

Based on this colouring one can then derive the following classical result (see Toselli and Widlund [128]):

THEOREM 5.2
The collection of subspaces $\{\mathcal{V}_i : i = 1, ..., p\}$ *can be coloured by* N_c *different colours so that when* $\mathcal{V}_i$ *and* $\mathcal{V}_j$, $i \neq j$, *have the same colour, we necessarily have* $\mathcal{V}_i$ *and* $\mathcal{V}_j$ *mutually orthogonal in the inner product introduced by the a-norm, and for the 1-level, resp. 2-level additive Schwarz preconditioner*

$$\lambda_{\max}\left(M_1^{-1}A\right) \leq N_c \quad \text{and} \quad \lambda_{\max}\left(M_{2,AS}^{-1}A\right) \leq N_c + 1. \tag{5.23}$$

For our problem the maximum eigenvalue of $M_1^{-1}A$, resp. $M_{2,AS}^{-1}A$ is always well behaved and for example bounded by 2 resp. 3 in the one dimensional and by 6 resp. 7 in the two dimensional case for a uniform mesh and a sufficiently small overlap (as shown in Figure 5-3).

For the minimum eigenvalues of our preconditioners we can use the following classical result:

THEOREM 5.3
Suppose, for each $l = 0, 1$, *there exists a constant* $C_l \geq 1$, *such that any function* $u_h \in \mathcal{V}^h$ *there exist functions* $u_j \in \mathcal{V}_j$, $j = l, \ldots, p$, *such that for the vectors corresponding to them*

$$\mathbf{u}_h = \sum_{j=l}^{p} \mathbf{u}_l \tag{5.24}$$

and

$$\sum_{j=l}^{p} a\left(\mathbf{u}_j, \mathbf{u}_j\right) \leq C_l a\left(\mathbf{u}_h, \mathbf{u}_h\right), \tag{5.25}$$

then

$$\lambda_{\min}\left(M_1^{-1}A\right) \geq 1/C_0 \quad \text{and} \quad \lambda_{\min}\left(M_{2,AS}^{-1}A\right) \geq 1/C_1. \tag{5.26}$$

Proof:
See for example in Chan and Mathew [28], page 93, Theorem 13. □

Combining Theorem 5.2 and Theorem 5.3 we finally get:

THEOREM 5.4
With N_c being the number of colours as defined above and with C_0 and C_1 being the constants from Theorem 5.3, the condition number $\kappa\left(M_1^{-1}A\right)$ of the one-level additive Schwarz preconditioner, resp. the condition number $\kappa\left(M_{2,AS}^{-1}A\right)$ of the two-level additive Schwarz preconditioner can be bounded by

$$\kappa\left(M_1^{-1}A\right) \leq N_c \cdot C_0, \quad \text{resp.} \quad \kappa\left(M_{2,AS}^{-1}A\right) \leq (N_c+1) \cdot C_1. \tag{5.27}$$

5.1.4 History of Schwarz methods

The first discussion of a method of decomposition of a domain into subdomains and of its convergence and was given by *Hermann Amandus Schwarz* already in 1870 (see Schwarz [118]). He developed his method as a way of solving boundary value problems on complicated domains. Not much attention was paid to his ideas for the following century, apart from perhaps the work of *Sergei Sobolev* in 1936 (see Sobolev [123]), who found the variational formulation of the Schwarz algorithm for an elasticity problem. The analysis was extended to finite element models and other methods derived from variation calculus. It took another 50 years until these ideas were revived in the 1980s. They were extended significantly since then and a large variety of literature on domain decomposition methods is now available.
Large parts of the theory on the *Schwarz alternating method* was developed by Pierre-Louis Lions. His work includes conditions for which the method converges for the preconditioned Darcy's equation, Stoke's problem, nonlinear monotone problems and evolution problems (see Lions [91]). These conditions are based on the maximum principle. He also generalized the method for the case of several subdomains (see Lions [92]) and for non-overlapping subdomains (see Lions [93]).
The importance of the *partition property* (as used in Theorem 5.3) was first discussed by Aleksandr M. Matsokin and Sergey V. Nepomnyashikh (see Matsokin and Nepomnyaschikh [99] and Nepomnyaschikh [106]).
Since the mid 1980s Maksymilian Dryja and Olof B. Widlund have contributed a lot to the *abstract theory for Additive Schwarz preconditioners* (see Dryja and Widlund [43]). They proved that the condition number of the preconditioned operator is uniformly bounded if the overlap between neighbouring subregions is large enough with respect to their actual size. Furthermore they managed to show that iterative substructuring methods with non-overlapping subdomains fit well into the general framework (Dryja and Widlund [42]) and pioneered the analysis and development of the preconditioners

in the 3D case (see Dryja [40, 41] and Dryja et al. [45]). Of even greater importance for this thesis however is their work on additive Schwarz algorithms with only small overlap (Dryja and Widlund [44]). The main advantage of small overlap clearly is the reduced cost of each iteration step. Dryja and Widlund derived among others the bound given in Lemma 5.13, which we use to find an improved dependence of the condition number of the preconditioned system on the overlap, replacing the proportionality of this condition number from $\left(1+\beta^{-2}\right)$ to $\left(1+\beta^{-1}\right)$. It is shown that the rate of convergence is quite satisfactory even in the case of minimum overlap. This theory was confirmed by a series of numerical results (see for example Cai [22, 23], Bjørstad et al. [13], Bjørstad and Skogen [12], Skogen [120] and Cai et al. [24]).

Next to refining the theory on additive Schwarz operators $P=\sum_{j=0}^{p} P_j$, other operators like the *multiplicative Schwarz operator* $P_{mul}=I-(I-P_0)(I-P_1)\cdots(I-P_p)$, where I is the identity operator (see for example Bramble et al. [16]), and *hybrid operators* like $P_{hyb}=P_0+(I-P_0)\left(\sum_{j=1}^{p} P_j\right)(I-P_0)$ (see Mandel [97]) were studied and a similar convergence theory was developed.

Next to these more theoretical discussions on Schwarz preconditioners more application based aspects were also considered. Due to their local solves, local storage possibilities and (in well-designed algorithms) limited need of communication between processors (especially additive) Schwarz preconditioners are well suited for parallel computers. Ways to improve their implementation in parallel were therefore discussed and numerical results on parallel machines given by Gropp [63], Gropp and Keyes [64, 65, 66], Bjørstad and Skogen [12], Bjørstad et al. [13], Bjørstad et al. [14], Chan [27] and Cliffe et al. [31].

Excellent, more *general overviews* of domain decomposition preconditioners and their history were given by Chan and Mathew [28] and more recently by Toselli and Widlund [128].

After recognizing these important contributions, we now would like to develop sharp condition number estimates of the additive Schwarz preconditioners $M_1^{-1}A(\alpha)$ and $M_{2,Lin}^{-1}A(\alpha)$ for problem (5.1,5.2). We want to find out more about, why these condition numbers can in practice sometimes be very large and how we can reduce them for two-level additive Schwarz preconditioners by modifying the coarsening operator.

5.2 Convergence estimates for the additive Schwarz preconditioned problem

While in the literature described above the dependence of the performance of additive Schwarz preconditioners on the fine mesh size h, the coarse mesh size H and the

minimum overlap βH was analysed in quite some detail, problems with large condition numbers caused by jumping coefficients α were mainly dealt with by using coarse meshes that resolved the jumps. As we will see in the numerical results at the end of this chapter, as well as in Chapters 6 and 7, the dependence on α is however very significant for groundwater flow problems as well as for other problems with strongly varying fields.
Resolving the jumps by the coarse grid for these problems however is problematic mainly due to the following two reasons:

- The coefficients vary strongly on a very small scale. Therefore we would need a very fine coarse grid to resolve this, which would increase the computational cost of applying the preconditioner significantly.
- We have to solve problem (5.1,5.2) for many realisations of the random field to be then able to apply some sort of Monte Carlo approach. Possible jumps however occur at different locations for different fields. Therefore a different coarse grid would have to be constructed for each field, which again increases the computational costs.

Let us now consider an additive Schwarz method with a coarse space of piecewise linear functions on the coarse grid. Let therefore for ψ_i^{Lin}, $i \in \{1, \dots n_H\}$, denote the piecewise linear coarse grid basis functions, i.e. for every coarse grid node $\mathbf{x}_k^H$, we have $\psi_i^{Lin}\left(\mathbf{x}_k^H\right) = \delta_{i,k}$ and ψ_i^{Lin} is piecewise linear on every coarse grid element $\bar{\Omega}_j \in \mathcal{T}^H(\bar{\Omega})$. Furthermore let $\boldsymbol{\psi}_1^{Lin}, \dots, \boldsymbol{\psi}_{n_H}^{Lin}$ be the vectors of values of the coarse grid basis functions at the fine grid nodes.

We then define in the case of coarse grid basis functions $R_{0,Lin}^T$ to be the $n_h \times n_H$ matrix

$$R_{0,Lin}^T := \left[\boldsymbol{\psi}_1^{Lin}, \dots, \boldsymbol{\psi}_{n_H}^{Lin}\right], \tag{5.28}$$

the *linear coarse grid stiffness matrix*

$$A_{0,Lin} := R_{0,Lin} A R_{0,Lin}^T. \tag{5.29}$$

and the *two level additive Schwarz preconditioner*

$$M_{2,Lin}^{-1} := M_1^{-1} + R_{0,Lin}^T A_{0,Lin}^{-1} R_{0,Lin}.$$

5.2.1 One level additive Schwarz

Based on Theorem 5.4 one can prove the following theorem.

THEOREM 5.5
There exists a positive constant C independent of h and H (but possibly dependent on α) such that for the one-level additive Schwarz preconditioned stiffness matrix

$$\kappa\left(M_1^{-1}A\right) \leq CH^{-2}\left(1+\beta^{-2}\right), \tag{5.30}$$

with minimum overlap βH as defined in the very beginning of section 5.1.2.

Proof:
See for example in Dryja and Widlund [42, 44]. □

As this bound shows, the one-level additive Schwarz preconditioner is not scalable, i.e. if we use it in an iterative method the convergence rate deteriorates, when we increase the number of subdomains, i.e. reduce H. This negative effect however can be removed by adding the second level, i.e. a coarse grid to our preconditioner. This is at an additional cost, which however is usually neglectable compared to the extra benefit for problems with many subdomains as the numerical results at the end of this chapter show.

5.2.2 Two level additive Schwarz

To derive some analytical results on the convergence rate of conjugate gradient methods with two-level additive Schwarz preconditioners we will need the following lemma.

LEMMA 5.6
There exists a family of functions $\chi_1, \ldots, \chi_p$ in $W^{1,\infty}(\Omega)$, subject to a covering $\mathcal{C}^{H,\beta} := \left\{\widetilde{\Omega_1}, \ldots, \widetilde{\Omega_p}\right\}$, such that

(i) $0 \leq \chi_j(\mathbf{x}) \leq 1, \quad \mathbf{x} \in \Omega, \quad j = 1, ..., p,$

(ii) $supp(\chi_j) \subset \widetilde{\Omega_j}, \quad j = 1, ..., p,$

(iii) $\sum_{j=1}^{p} \chi_j(\mathbf{x}) = 1, \quad \mathbf{x} \in \Omega,$

(iv) $|\nabla \chi_j(\mathbf{x})| \leq C\beta^{-1}H^{-1}, \quad \mathbf{x} \in \Omega, \quad j = 1, ..., p,$

for some constant C, independent of β and H.
This family of functions is then called a ***partition of unity****.*

Proof:
See Toselli and Widlund [128], pages 57-59, Lemma 3.4. □

Before we will deal with problems arising from strongly varying fields, we now first of all want to find out more about the dependence of $M_{2,Lin}^{-1}A(\alpha)$ on α. This was not done before for fields that are allowed to vary strongly within the subdomains. Let us therefore first of all prove the following new theorem.

THEOREM 5.7
For any $u_h \in \mathcal{V}^h$ there exist $u_i \in \mathcal{V}_i$, $i = 0, ..., p$, such that

$$u_h = \sum_{i=0}^{p} u_i, \tag{5.31}$$

and

$$\sum_{i=0}^{p} |u_i|^2_{H^1(\Omega),\alpha} \leq C\left(1 + \beta^{-2}\right)\left(\max_j \max_{\mathbf{x},\mathbf{y}\in\Omega_j}\left(\frac{\alpha(\mathbf{x})}{\alpha(\mathbf{y})}\right)\right) B(d)\, |u_h|^2_{H^1(\Omega),\alpha}, \tag{5.32}$$

where $B(1) = 1$, $B(2) = (1 + \log(H/h))$ and $B(3) = (H/h)$.

Proof:
Recall that Ω_j, $j = 1, ..., p$, are the non-overlapping subdomains. We then know that for any $v_h \in \mathcal{V}_j$, we have

$$|v_h|^2_{H^1(\Omega_j)} = \int_{\Omega_j} \nabla v_h . \nabla v_h \quad \text{and} \quad |v_h|^2_{H^1(\Omega_j),\alpha} = \int_{\Omega_j} \alpha \nabla v_h . \nabla v_h. \tag{5.33}$$

As $\alpha(\mathbf{x}) > 0$ for all $\mathbf{x} \in \Omega_j$, we get

$$\min_{\mathbf{x}\in\Omega_j}(\alpha(\mathbf{x})) \cdot |v_h|^2_{H^1(\Omega_j),\alpha} \leq |v_h|^2_{H^1(\Omega_j),\alpha} \leq \max_{\mathbf{x}\in\Omega_j}(\alpha(\mathbf{x})) \cdot |v_h|^2_{H^1(\Omega_j)}. \tag{5.34}$$

Now let $\mathcal{C}(\bar{\Omega})$ be the space of continuous functions on the closure $\bar{\Omega}$ of Ω. Then define $I^H : \mathcal{C}(\bar{\Omega}) \leftarrow \mathcal{V}_0$ to be the interpolation operator, such that for $u_h \in \mathcal{C}(\bar{\Omega})$, $u_0 := I^H u_h$ is the result of piecewise linear interpolation of u_h onto the coarse space $\mathcal{V}_0$, i.e. it projects the values of u_h at the coarse grid nodes onto themselves and interpolates linearly between these values on the coarse grid elements.

Then as shown for example by Toselli and Widlund [128], page 100, there exists a generic constant C, depending on the dimension of the problem, but independent of α, h and H, such that in one dimension

$$|u_0|^2_{H^1(\Omega_j)} \leq C\,|u_h|^2_{H^1(\Omega_j)}\,, \tag{5.35}$$

in two dimensions

$$|u_0|^2_{H^1(\Omega_j)} \leq C\,(1+\log(H/h))\,|u_h|^2_{H^1(\Omega_j)} \tag{5.36}$$

and in three dimensions

$$|u_0|^2_{H^1(\Omega_j)} \leq C\,(H/h)\,|u_h|^2_{H^1(\Omega_j)}\,. \tag{5.37}$$

Combining (5.34) with these bounds, we get

$$|u_0|^2_{H^1(\Omega_j),\alpha} \leq C \cdot \left(\max_{\mathbf{x},\mathbf{y}\in\Omega_j} \frac{\alpha(\mathbf{x})}{\alpha(\mathbf{y})}\right) \cdot B(d) \cdot |u_h|^2_{H^1(\Omega_j),\alpha}\,, \tag{5.38}$$

where

$$B(d) := \begin{cases} 1, & \text{if } d=1, \\ (1+\log(H/h))\,, & \text{if } d=2, \\ (H/h)\,, & \text{if } d=3. \end{cases}$$

Applying the standard finite element interpolation error bound for $(u_h - u_0)$ (see again for example in Toselli and Widlund [128], page 100), we obtain

$$\|u_h - u_0\|^2_{L^2(\Omega_j)} \leq C \cdot B(d) \cdot H^2\,|u_h|^2_{H^1(\Omega_j)}\,. \tag{5.39}$$

Choosing a partition of unity as defined in Definition 5.6, let us now define the following partition of $u_h(\mathbf{x}) - u_0(\mathbf{x})$, $\forall \mathbf{x} \in \Omega$:

$$u_i(\mathbf{x}) := I^h\left(\chi_i\,(u_h - u_0)\,(\mathbf{x})\right), \qquad \text{for } i = 1,\ldots,p, \tag{5.40}$$

where I^h is the finite element interpolation onto $\mathcal{V}^h$, which is needed here, as $\chi_i (u_h - u_0)$ is not automatically in $\mathcal{V}^h$.

Now let τ be an element in $\bar{\Omega}_j$ and let $\hat{\chi}_{i,\tau}$ be the average of χ_i over τ. Since $\mathcal{T}^h(\bar{\Omega})$ is a refinement of $\mathcal{T}^H(\bar{\Omega})$ we see that $(u_h - u_0) \in \mathcal{V}^h$. Therefore $I^h (u_h - u_0) = u_h - u_0$. We get for any $\mathbf{x} \in \tau$

$$\begin{aligned} u_i(\mathbf{x}) &= I^h \left(\chi_i \left(u_h - u_0\right)(\mathbf{x})\right) \\ &= I^h \left(\left(\chi_i - \hat{\chi}_{i,\tau}\right)\left(u_h - u_0\right)(\mathbf{x})\right) + I^h \left(\hat{\chi}_{i,\tau}\left(u_h - u_0\right)(\mathbf{x})\right) \\ &= I^h \left(\left(\chi_i - \hat{\chi}_{i,\tau}\right)\left(u_h - u_0\right)(\mathbf{x})\right) + \hat{\chi}_{i,\tau}\left(u_h - u_0\right)(\mathbf{x}) \end{aligned} \tag{5.41}$$

and therefore

$$\boldsymbol{\nabla} u_i(\mathbf{x}) = \boldsymbol{\nabla} I^h \left(\left(\chi_i - \hat{\chi}_{i,\tau}\right)\left(u_h - u_0\right)(\mathbf{x})\right) + \boldsymbol{\nabla}\left(\hat{\chi}_{i,\tau}\left(u_h(\mathbf{x}) - u_0(\mathbf{x})\right)(\mathbf{x})\right). \tag{5.42}$$

Integrating over τ and applying the triangle inequality to the resulting equation, we obtain

$$|u_i|^2_{H^1(\tau)} \leq 2 \left| I^h \left(\left(\chi_i - \hat{\chi}_{i,\tau}\right)\left(u_h - u_0\right)\right)\right|^2_{H^1(\tau)} + 2\left|\hat{\chi}_{i,\tau}\left(u_h - u_0\right)\right|^2_{H^1(\tau)}. \tag{5.43}$$

We can now apply an inverse inequality (i.e. $\forall v_h \in \mathcal{V}^h \quad |v_h|_{H^1(\tau)} \leq Ch^{-1} \|v_h\|_{L_2(\tau)}$), the fact, that $\left\|I^h (f \cdot v_h)\right\|_{L_2(\tau)} \leq \|f\|_{L_\infty(\tau)} \|v_h\|_{L_2(\tau)}$ for any continuous function f, and the fact, that χ_i is continuous with $|\hat{\chi}_{i,\tau}(\mathbf{x})| \leq 1$ to get

$$|u_i|^2_{H^1(\tau)} \leq Ch^{-2} \|\chi_i - \hat{\chi}_{i,\tau}\|^2_{L_\infty(\tau)} \|u_h - u_0\|^2_{L_2(\tau)} + 2\left|u_h - u_0\right|^2_{H^1(\tau)}. \tag{5.44}$$

By property (iv) of the partition of unity, we get

$$\|\chi_i - \hat{\chi}_{i,\tau}\|_{L_\infty(\tau)} \leq C\beta^{-1} H^{-1} h. \tag{5.45}$$

Using this inequality in (5.44), we obtain

$$\begin{aligned} |u_i|^2_{H^1(\tau)} &\leq C\left(h^{-2}\left(\frac{h}{\beta H}\right)^2 \|u_h - u_0\|^2_{L_2(\tau)} + |u_h - u_0|^2_{H^1(\tau)}\right) \\ &\leq C\left(\beta^{-2} H^{-2} \|u_h - u_0\|^2_{L_2(\tau)} + |u_h - u_0|^2_{H^1(\tau)}\right). \end{aligned} \tag{5.46}$$

Summing over all $i \in \{1, \ldots, p\}$ and noting that only a finite number of u_i, bounded by the minimum number of colourings, N_c, from the colouring argument (see subsection

5.1.3), is nonzero on any fine grid element τ, we obtain

$$\sum_{i=1}^{p} |u_i|^2_{H^1(\tau)} \leq C \left(\beta^{-2} H^{-2} \|u_h - u_0\|^2_{L_2(\tau)} + |u_h - u_0|^2_{H^1(\tau)} \right) N_c. \tag{5.47}$$

Now summing over all elements $\tau \subset \bar{\Omega}_j$ and combining all constants, i.e. C and N_c, in a generalised constant C, we get for all $j \in \{1, ..., p\}$

$$\sum_{i=1}^{p} |u_i|^2_{H^1(\Omega_j)} \leq C \left(\beta^{-2} H^{-2} \|u_h - u_0\|^2_{L_2(\Omega_j)} + |u_h - u_0|^2_{H^1(\Omega_j)} \right). \tag{5.48}$$

Applying the standard interpolation bound (5.39) to the first term on the right and using the fact that $(a-b)^2 \leq 2a^2 + 2b^2$ for any $a, b \in \mathbb{R}$ for the second term on the right, we get

$$\begin{aligned} \sum_{i=1}^{p} |u_i|^2_{H^1(\Omega_j)} &\leq C \left(\beta^{-2} H^{-2} B(d) H^2 |u_h|^2_{H^1(\Omega_j)} + 2 |u_h|^2_{H^1(\Omega_j)} + 2 |u_0|^2_{H^1(\Omega_j)} \right) \\ &\leq C \left(\beta^{-2} B(d) |u_h|^2_{H^1(\Omega_j)} + |u_h|^2_{H^1(\Omega_j)} + B(d) |u_h|^2_{H^1(\Omega_j)} \right). \end{aligned} \tag{5.49}$$

Applying inequalities (5.34), summing over all subdomains and adding (5.38) to each side, we obtain

$$\sum_{i=0}^{p} |u_i|^2_{H^1(\Omega),\alpha} \leq C \cdot (1 + \beta^{-2}) \left(\max_j \max_{\mathbf{x},\mathbf{y} \in \bar{\Omega}_j} \left(\frac{\alpha(\mathbf{x})}{\alpha(\mathbf{y})} \right) \right) B(d) |u_h|^2_{H^1(\Omega),\alpha}, \tag{5.50}$$

which finishes the proof.

□

The bound given in Theorem 5.7 is independent of coefficient jumps between different subdomains, which shows that resolving the coefficients $\boldsymbol{\alpha}$ by the coarse grid leads to two-level additive Schwarz preconditioners $M^{-1}_{2,Lin}$ that are independent of the underlying field, i.e. we then get $\kappa \left(M^{-1}_{2,Lin} A(\boldsymbol{\alpha}) \right) = \kappa \left(M^{-1}_{2,Lin} A(\mathbf{1}) \right)$. However the dependence on the ratio H/h in three dimensional problems is quite strong and it often might be desirable to get rid of this dependence, possibly paying for it by allowing the ratio of coefficients being taken over larger regions and not only within subdomains.

Let us therefore (for theoretical purposes only) define a quasi-interpolant $\bar{I}^H : H^1_0(\Omega) \to \mathcal{V}_0$ as follows. Let $\mathbf{x}^H_k$ be a node of the coarse mesh $\mathcal{T}^H(\bar{\Omega})$ and let ω_k be the union of

coarse grid elements, that share $\mathbf{x}_k^H$. For coarse grid node $\mathbf{x}_k^H$, the *quasi-interpolant* is now defined to map $u \in H_0^1(\Omega)$ to $\mathcal{V}_0$ and to satisfy

$$\left(\bar{I}^H u\right)\left(\mathbf{x}_k^H\right) := \begin{cases} 0, & \mathbf{x}_k^H \in \Gamma, \\ |\omega_k|^{-1} \int\limits_{\omega_k} u(\mathbf{x}) \mathrm{d}x, & \text{otherwise}, \end{cases} \tag{5.51}$$

where $|\omega_k|$ denotes the measure of ω_k.

Given an element $\bar{\Omega}_j \in \mathcal{T}^H(\bar{\Omega})$, we define ω_{Ω_j} as the union of Ω_j and its neighbouring subdomains (see Figure 5-4 and as defined before in 5.16)

$$\omega_{\Omega_j} := \text{interior}\left(\bigcup_{k:\bar{\Omega}_k \in \mathcal{T}^H(\bar{\Omega}), \bar{\Omega}_k \cap \bar{\Omega}_j \neq \emptyset} \bar{\Omega}_k\right) \tag{5.52}$$

and $\bar{\omega}_{\Omega_j}$ to be its closure.

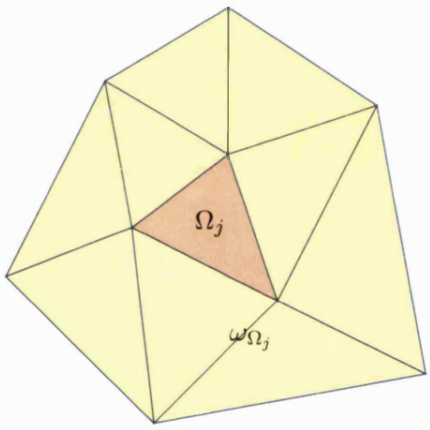


Figure 5-4: Subdomain with its neighbouring subdomains

The following result is proved in Toselli and Widlund [128], page 63, Lemma 3.6:

LEMMA 5.8
Let $\mathcal{T}^H(\Omega)$ be shape regular and $u \in H_0^1(\Omega)$. Then there exists a (generic) constant C (depending on the dimension of the problem, but not on α, h or H), such that

$$\left\|u - \bar{I}^H u\right\|_{L_2(\Omega_j)} \leq C \cdot H \cdot |u|_{H^1\left(\omega_{\Omega_j}\right)}, \tag{5.53}$$

$$\left|\bar{I}^H u\right|_{H^1(\Omega_j)} \leq C \cdot |u|_{H^1\left(\omega_{\Omega_j}\right)}. \tag{5.54}$$

Using this Lemma we can prove the following new theorem.

THEOREM 5.9
For any $u_h \in \mathcal{V}^h$ there exist $u_i \in \mathcal{V}_i$, $i = 0, ..., p$, such that

$$u_h := \sum_{i=0}^{p} u_i \tag{5.55}$$

and

$$\sum_{i=0}^{p} |u_i|^2_{H^1(\Omega),\alpha} \leq C\left(1+\beta^{-2}\right)\left(\max_j \max_{\mathbf{x},\mathbf{y}\in\omega_{\Omega_j}}\left(\frac{\alpha(\mathbf{x})}{\alpha(\mathbf{y})}\right)\right)|u_h|^2_{H^1(\Omega),\alpha}. \tag{5.56}$$

Proof:
For any $v_h \in \mathcal{V}^h$ with

$$|v_h|^2_{H^1\left(\omega_{\Omega_j}\right)} = \int_{\omega_{\Omega_j}} \nabla v_h . \nabla v_h \quad \text{and} \quad |v_h|^2_{H^1\left(\omega_{\Omega_j}\right),\alpha} = \int_{\omega_{\Omega_j}} \alpha \nabla v_h . \nabla v_h. \tag{5.57}$$

Since as $\alpha(\mathbf{x}) > 0$ for all $\mathbf{x} \in \omega_{\Omega_j}$ we have

$$\min_{\mathbf{x}\in\omega_{\Omega_j}}(\alpha(\mathbf{x}))\cdot|v_h|^2_{H^1\left(\omega_{\Omega_j}\right)} \leq |v_h|^2_{H^1\left(\omega_{\Omega_j}\right),\alpha} \leq \max_{\mathbf{x}\in\omega_{\Omega_j}}(\alpha(\mathbf{x}))\cdot|v_h|^2_{H^1\left(\omega_{\Omega_j}\right)}. \tag{5.58}$$

Now let $u_0 := \bar{I}^H u_h$. Then we have by (5.54) independent of the dimension of the problem

$$|u_0|^2_{H^1(\Omega_j)} \leq C\,|u_h|^2_{H^1\left(\omega_{\Omega_j}\right)}. \tag{5.59}$$

Combining inequalities (5.58) and (5.59), we get

$$|u_0|^2_{H^1\left(\omega_{\Omega_j}\right),\alpha} \leq C \cdot \left(\max_{\mathbf{x},\mathbf{y}\in\omega_{\Omega_j}}\left(\frac{\alpha(\mathbf{x})}{\alpha(\mathbf{y})}\right)\right)\cdot|u_h|^2_{H^1\left(\omega_{\Omega_j}\right),\alpha}. \tag{5.60}$$

Choosing a fine grid element τ in the closure of ω_{Ω_j} the next steps of the proof are identical to steps (5.39) to (5.46) in the proof of Theorem 5.7 to obtain the inequality

$$|u_i|^2_{H^1(\tau)} \leq C\left(\beta^{-2}H^{-2}\left\|u_h - u_0\right\|^2_{L_2(\tau)} + |u_h - u_0|^2_{H^1(\tau)}\right). \tag{5.61}$$

Summing over all $i \in \{1, \ldots, p\}$ and noting that only a finite number of u_i, bounded by the minimum of colourings, N_c, from the colouring argument, is nonzero on any finite grid element τ, we obtain

$$\sum_{i=1}^{p} |u_i|^2_{H^1(\tau)} \leq C\left(\beta^{-2}H^{-2}\left\|u_h - u_0\right\|^2_{L_2(\tau)} + |u_h - u_0|^2_{H^1(\tau)}\right) N_c. \tag{5.62}$$

We sum over all $\tau \subset \bar{\Omega}_j$ and combine C and N_c in a generalised constant C, to get for all subdomains Ω_j

$$\sum_{i=1}^{p} |u_i|^2_{H^1(\Omega_j)} \leq C\left(\beta^{-2}H^{-2}\left\|u_h - u_0\right\|^2_{L_2(\Omega_j)} + |u_h - u_0|^2_{H^1(\Omega_j)}\right). \tag{5.63}$$

Applying inequality (5.53) to the first term on the right and the triangle inequality and (5.54) to the second term on the right, we get

$$\begin{aligned}
\sum_{i=1}^{p} |u_i|^2_{H^1(\Omega_j)} &\leq C\left(\beta^{-2}H^{-2}H^2\,|u_h|^2_{H^1\left(\omega_{\Omega_j}\right)} + |u_h|^2_{H^1(\Omega_j)} + |u_0|^2_{H^1(\Omega_j)}\right) \\
&\leq C\left(\beta^{-2}\,|u_h|^2_{H^1\left(\omega_{\Omega_j}\right)} + |u_h|^2_{H^1(\Omega_j)} + |u_h|^2_{H^1(\Omega_j)}\right) \\
&\leq C\left(1+\beta^{-2}\right)|u_h|^2_{H^1\left(\omega_{\Omega_j}\right)}.
\end{aligned} \tag{5.64}$$

Applying inequality (5.58), summing over all subregions Ω_j and adding (5.59) to both sides of the inequality, we obtain

$$\sum_{i=0}^{p} |u_i|^2_{H^1(\Omega),\alpha} \leq C\left(1+\beta^{-2}\right)\left(\max_j \max_{\mathbf{x},\mathbf{y}\in\omega_{\Omega_j}}\left(\frac{\alpha(\mathbf{x})}{\alpha(\mathbf{y})}\right)\right)|u_h|^2_{H^1(\Omega),\alpha}, \tag{5.65}$$

which finishes the proof.

□

The bounds in Theorem 5.7 and 5.9 can be improved further, replacing the factor $\left(1+\beta^{-2}\right)$ by a factor $\left(1+\beta^{-1}\right)$. This we shall do in Lemma 5.13 and Theorem 5.16

following an idea first developed by Dryja and Widlund [44].

The next theorem (a proof of which can be found in Nečas [103], Chapter 7) will be quoted here without proof.

THEOREM 5.10
Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain and let f_k, $k = 1, ..., L$, $L \geq 1$, be (not necessarily linear) functionals in $H^1(\Omega)$, such that if v is constant on Ω

$$\sum_{k=1}^{L} |f_k(v)|^2 = 0 \quad \Leftrightarrow \quad v = 0.$$

Then there exist constants C_1 and C_2, depending only on Ω and the functionals f_i, such that for $u \in H^1(\Omega)$,

$$\|u\|^2_{L^2(\Omega)} \leq C_1 |u|^2_{H^1(\Omega)} + C_2 \sum_{k=1}^{L} |f_k(u)|^2 . \tag{5.66}$$

Let $\Gamma \subseteq \bar{\Omega}$ have (d-1)-dimensional measure, then applying Theorem 5.10 with $L = 1$ and $f_1(u) = \|u\|_{L^2(\Gamma)}$ results directly in the following Lemma.

LEMMA 5.11
***(Friedrich's inequality)** There exist constants C_1 and C_2, depending only on Ω and Γ, such that for $u \in H^1(\Omega)$,*

$$\|u\|^2_{L^2(\Omega)} \leq C_1 |u|^2_{H^1(\Omega)} + C_2 \|u\|^2_{L^2(\Gamma)} . \tag{5.67}$$

In particular, if u vanishes on Γ,

$$\|u\|^2_{L^2(\Omega)} \leq C_1 |u|^2_{H^1(\Omega)}$$

and thus

$$|u|^2_{H^1(\Omega)} \leq \|u\|^2_{H^1(\Omega)} \leq (C_1 + 1) |u|^2_{H^1(\Omega)} .$$

DEFINITION 5.12
Let $\partial\Omega_j$, $j = 1, ..., p$, denote the boundary of subdomain Ω_j and $\Gamma_j := \partial\Omega_j \backslash \Gamma$. Furthermore let $\Gamma_{\beta,j} \subset \Omega_j$ denote the set of points within a distance βH of Γ_j and let $\bar{\Gamma}_{\beta,j}$ denote its closure.

Closely related to Friedrich's inequality, we can then prove the following Lemma.

LEMMA 5.13
For any $u \in H^1(\Omega_j)$ *we get*

$$\|u\|^2_{L_2(\Gamma_{\beta,j})} \leq C\beta H^2 (1+\beta) |u|^2_{H^1(\Omega_j)} + C\beta \|u\|^2_{L_2(\Omega_j)}. \tag{5.68}$$

Proof:
The proof of this Lemma is based on Theorem 5.10 and Lemma 5.11 and is given in Toselli and Widlund [128], Lemma 3.10. □

Based on this we can then prove the following new result.

THEOREM 5.14
For any $u_h \in \mathcal{V}^h$ *there exist* $u_i \in \mathcal{V}_i$, $i = 0, ..., p$, *such that*

$$u_h = \sum_{i=0}^{p} u_i, \tag{5.69}$$

and

$$\sum_{i=0}^{p} |u_i|^2_{H^1(\Omega),\alpha} \leq C\left(1+\beta^{-1}\right) \left(\max_j \max_{\mathbf{x},\mathbf{y}\in\Omega_j} \left(\frac{\alpha(\mathbf{x})}{\alpha(\mathbf{y})} \right) \right) B(d) |u_h|^2_{H^1(\Omega),\alpha}, \tag{5.70}$$

where $B(1) = 1$, $B(2) = (1 + \log(H/h))$ *and* $B(3) = (H/h)$.

Proof:
The first part of this proof is identical to that for Theorem 5.7 up to inequality (5.46). By the properties of the partition of unity (see Definition 5.6) and the definition of $\Gamma_{\beta,j}$ (see Definition 5.12) it follows for any $i = 1, ..., p$, that $\chi_i(\mathbf{x}) = \delta_{ij}$ for $\mathbf{x} \in \Omega_j \backslash \Gamma_{\beta,j}$, $j = 1, ..., p$. Therefore with $u_i := I^h(\chi_i(u_h - u_0))$, where I^h is the finite element

interpolation onto $\mathcal{V}^h$, we obtain

$$\begin{aligned}\sum_{i=1}^{p} |u_i|^2_{H^1(\Omega_j\backslash\Gamma_{\beta,j})} &= |u_j|^2_{H^1(\Omega_j\backslash\Gamma_{\beta,j})} \\ &= \left|I^h\left(\chi_j(u_h-u_0)\right)\right|^2_{H^1(\Omega_j\backslash\Gamma_{\beta,j})} \\ &= |u_h-u_0|^2_{H^1(\Omega_j\backslash\Gamma_{\beta,j})},\end{aligned} \tag{5.71}$$

since $\chi_j(\mathbf{x}) = 1$ for $\mathbf{x} \in \Omega_j\backslash\Gamma_{\beta,j}$ and since $u_h - u_0$ is piecewise linear on the whole of Ω with respect to the fine grid.

Now consider a fine grid element $\tau \subset \bar{\Gamma}_{\beta,j}$.
We then get from (5.46)

$$\sum_{i=1}^{p} |u_i|^2_{H^1(\tau)} \leq C\left(\beta^{-2}H^{-2}\|u_h-u_0\|^2_{L_2(\tau)} + |u_h-u_0|^2_{H^1(\tau)}\right) N_c. \tag{5.72}$$

Summing (5.72) over all fine grid elements $\tau \subset \bar{\Gamma}_{\beta,j}$ and combining all constants, i.e. C and N_c in a generalized constant C, we get

$$\sum_{i=1}^{p} |u_i|^2_{H^1(\Gamma_{\beta,j})} \leq C\beta^{-2}H^{-2}\|u_h-u_0\|^2_{L_2(\Gamma_{\beta,j})} + C\,|u_h-u_0|^2_{H^1(\Gamma_{\beta,j})}. \tag{5.73}$$

To estimate $\|u_h - u_0\|_{L_2(\Gamma_{\beta,j})}$ we now use Lemma 5.13 to obtain

$$\|u_h-u_0\|^2_{L_2(\Gamma_{\beta,j})} \leq C\beta H^2(1+\beta)\,|u_h-u_0|^2_{H^1(\Omega_j)} + C\beta\,\|u_h-u_0\|^2_{L_2(\Omega_j)}. \tag{5.74}$$

Employing (5.74) in (5.73) we get

$$\sum_{i=1}^{p} |u_i|^2_{H^1(\Gamma_{\beta,j})} \leq C\left(1+\beta^{-1}\right)|u_h-u_0|^2_{H^1(\Omega_j)} + C\beta^{-1}H^{-2}\|u_h-u_0\|^2_{L_2(\Omega_j)}. \tag{5.75}$$

Adding (5.71) to (5.75), we then obtain

$$\sum_{i=1}^{p} |u_i|^2_{H^1(\Omega_j)} \leq C\left(1+\beta^{-1}\right)|u_h-u_0|^2_{H^1(\Omega_j)} + C\beta^{-1}H^{-2}\|u_h-u_0\|^2_{L_2(\Omega_j)}. \tag{5.76}$$

Applying the triangle inequality to the first term on the right and the standard interpolation error inequality (5.39) to the second term on the right (as it was done in the

step from (5.48) to (5.49)), we get

$$\sum_{i=1}^{p} |u_i|^2_{H^1(\Omega_j)} \leq C\left(1+\beta^{-1}\right) B(d)\, |u_h|^2_{H^1(\Omega_j)} + C\left(1+\beta^{-1}\right) \|u_0\|^2_{H^1(\Omega_j)} + C\beta^{-1}\, |u_h|^2_{H^1(\Omega_j)}\,, \tag{5.77}$$

with $B(d)$ as defined above.
Using (5.38) and combining terms using (5.35) (in 1D), (5.36) (in 2D), resp. (5.37) (in 3D), we obtain

$$\sum_{i=1}^{p} |u_i|^2_{H^1(\Omega_j)} \leq C\left(1+\beta^{-1}\right) B(d)\, |u_h|^2_{H^1(\Omega_j)}\,. \tag{5.78}$$

Applying

$$\min_{\mathbf{x}\in\Omega_j} \left(\alpha\left(\mathbf{x}\right)\right) \cdot |u_h|^2_{H^1(\Omega_j)} \leq |u_h|^2_{H^1(\Omega_j),\alpha} \leq \max_{\mathbf{x}\in\Omega_j} \left(\alpha\left(\mathbf{x}\right)\right) \cdot |u_h|^2_{H^1(\Omega_j)}\,, \tag{5.79}$$

adding (5.38) on both sides of the inequality and summing over all subregions, it follows that

$$\sum_{i=0}^{p} |u_i|^2_{H^1(\Omega),\alpha} \leq C\left(1+\beta^{-1}\right) B(d) \left(\max_j \max_{\mathbf{x},\mathbf{y}\in\Omega_j} \left(\frac{\alpha\left(\mathbf{x}\right)}{\alpha\left(\mathbf{y}\right)}\right)\right) |u_h|^2_{H^1(\Omega),\alpha}\,, \tag{5.80}$$

which finishes the proof. □

Similarly to Theorem 5.14 we can improve the bound given in Theorem 5.9 to obtain the following new bound.

THEOREM 5.15
For any $u_h \in \mathcal{V}^h$ there exist $u_i \in \mathcal{V}_i$, $i = 0, ..., p$, such that

$$u_h := \sum_{i=0}^{p} u_i \tag{5.81}$$

and

$$\sum_{i=0}^{p} |u_i|^2_{H^1(\Omega),\alpha} \leq C\left(1+\beta^{-1}\right) \left(\max_j \max_{\mathbf{x},\mathbf{y}\in\omega_{\Omega_j}} \left(\frac{\alpha\left(\mathbf{x}\right)}{\alpha\left(\mathbf{y}\right)}\right)\right) |u_h|^2_{H^1(\Omega),\alpha}\,, \tag{5.82}$$

Proof:
Exactly as in Theorem 5.14, except that the interpolation I^H is now replaced by a

quasi-interpolant $\bar{I}^H$ (see Theorem 5.9). □

Using Theorem 5.4 we now obtain the following new result for the condition number of the preconditioned system:

THEOREM 5.16
There exists a constant C independent of α, h and H, such that the condition number $\kappa\left(M_{2,Lin}^{-1}A\right)$ of the two-level additive Schwarz preconditioner can be bounded by

$$\kappa\left(M_{2,Lin}^{-1}A\right) \leq C\left(1+\beta^{-1}\right)\left(\max_j \max_{\mathbf{x},\mathbf{y}\in\Omega_j}\left(\frac{\alpha(\mathbf{x})}{\alpha(\mathbf{y})}\right)\right) B(d), \tag{5.83}$$

where $B(1)=1$, $B(2)=(1+\log(H/h))$ and $B(3)=(H/h)$.

Combining Theorem 5.4 with Theorem 5.15, we get the new bound:

THEOREM 5.17
There exists a constant C independent of α, h and H, such that the condition number $\kappa\left(M_{2,Lin}^{-1}A\right)$ of the two-level additive Schwarz preconditioner can be bounded by

$$\kappa\left(M_{2,Lin}^{-1}A\right) \leq C\left(1+\beta^{-1}\right)\left(\max_j \max_{\mathbf{x},\mathbf{y}\in\omega_{\Omega_j}}\left(\frac{\alpha(\mathbf{x})}{\alpha(\mathbf{y})}\right)\right). \tag{5.84}$$

This finishes our theoretical framework of bounds for $\kappa\left(M_{2,Lin}^{-1}A\right)$. While Theorem 5.16 gives a bound that only depends on ratios of the coefficient α within one subdomain for the cost of depending on $\log(H/h)$ in 2D and even on H/h in 3D, Theorem 5.17 gets rid of this dependence on the ratio of the mesh sizes by paying for this by depending on ratios of α in the neighbourhood ω_{Ω_j} of a subdomain Ω_j. The bound in Theorem 5.16 is therefore more precise in the case of strong jumps of α between the subdomains, while the bound in Theorem 5.17 can be used, if big jumps happen mainly or also inside single subdomains. While ideas used in the proof of Theorem 5.16 were discussed before (e.g. by Chan and Mathew [28]), bounds as in Theorem 5.17 seem to be new and no suitable reference was found by the author.

5.3 Probability estimates

With these condition number bounds, we now try to find expectation estimates and probability bounds for the condition numbers for the preconditioned matrix (5.20) with underlying permeability field α on Ω. With the theory given in Chapters 3 and 4 we will give bounds on the expectation numbers as well as probability distributions for fields of Ornstein-Uhlenbeck type with correlation length λ.

As shown in the previous section the sharp upper bounds of the condition numbers of the two level additive Schwarz preconditioned systems with linear coarsening depends linearly on the ratio $\max_j \max_{\mathbf{x},\mathbf{y}\in\Omega_j} \left(\frac{\alpha(\mathbf{x})}{\alpha(\mathbf{y})}\right)$ (needed for the bound in Theorem 5.16) or on the ratio $\max_j \max_{\mathbf{x},\mathbf{y}\in\omega_{\Omega_j}} \left(\frac{\alpha(\mathbf{x})}{\alpha(\mathbf{y})}\right)$ (needed for the bound in Theorem 5.17).

Assuming constant permeability α^τ on each fine grid element τ the problem reduces to finding expectation estimates for $\max_j \max_{\tau,\tau'\in\mathcal{T}^h(\bar{\Omega}_j)} \left(\frac{\alpha^\tau}{\alpha^{\tau'}}\right)$ and $\max_j \max_{\tau,\tau'\in\mathcal{T}^h(\bar{\omega}_{\Omega_j})} \left(\frac{\alpha^\tau}{\alpha^{\tau'}}\right)$.

Let us now focus again on lognormal random fields with covariance of Ornstein-Uhlenbeck type (see Section 3.1) with correlation length λ.

5.3.1 Expectation estimates based on simple norm inequalities

Let us assume in the following, that the mesh $\mathcal{T}^h(\bar{\Omega})$ is resolved fine enough such that for every $\tau \in \mathcal{T}^h(\bar{\Omega})$, $\alpha^\tau := \frac{1}{|\tau|}\int_\tau \alpha$ can be approximated well enough and replaced by setting $\alpha^\tau := \alpha(\mathbf{m}^\tau)$, where $\mathbf{m}^\tau$ is the centroid of τ. Denoting $\#\mathcal{T}^h_{\Omega_j}$ for the number of fine grid elements τ in $\bar{\Omega}_j$ and p for the number of subdomains in Ω, we get the (often quite pessimistic) bound

LEMMA 5.18
If H is the maximum diameter of the subdomains Ω_j, $j = 1,...,p$, then

$$\mathbb{E}\left[\max_j \max_{\tau,\tau'\in\mathcal{T}^h(\bar{\Omega}_j)} \left(\frac{\alpha^\tau}{\alpha^{\tau'}}\right)\right] \leq p \cdot \left(\#\mathcal{T}^h_{\Omega_j}\right)^2 \cdot \exp\left\{\sigma^2\left\{1-\exp\left\{-H/\lambda\right\}\right\}\right\}. \qquad (5.85)$$

Proof:
Let $D := p \cdot \left(\#\mathcal{T}^h_{\bar{\Omega}_j}\right)^2$ and let $\mathbf{b} \in \mathbb{R}^D$ be the vector with entries

$$b_{j,\tau,\tau'} := \left(\frac{\alpha^\tau}{\alpha^{\tau'}}\right), \text{ for } \tau, \tau' \in \mathcal{T}^h(\bar{\Omega}_j). \tag{5.86}$$

Proceeding in a similar way as in Chapter 4 we get

$$\begin{aligned}
\mathbb{E}\left[\max_j \max_{\tau,\tau' \in \mathcal{T}^h(\bar{\Omega}_j)} \left(\frac{\alpha^\tau}{\alpha^{\tau'}}\right)\right] &= \mathbb{E}\left[\max_j \max_{\tau,\tau' \in \mathcal{T}^h(\bar{\Omega}_j)} (\mathbf{b})\right] \\
&= \mathbb{E}\left[\|\mathbf{b}\|_\infty\right] \\
&\leq \mathbb{E}\left[\|\mathbf{b}\|_1\right] \\
&= \mathbb{E}\left[\sum_j \sum_{\tau \in \mathcal{T}^h(\bar{\Omega}_j)} \sum_{\tau' \in \mathcal{T}^h(\bar{\Omega}_j)} \left(\frac{\alpha^\tau}{\alpha^{\tau'}}\right)\right] \\
&= \sum_j \sum_{\tau \in \mathcal{T}^h(\bar{\Omega}_j)} \sum_{\tau' \in \mathcal{T}^h(\bar{\Omega}_j)} \mathbb{E}\left[\frac{\alpha^\tau}{\alpha^{\tau'}}\right].
\end{aligned} \tag{5.87}$$

Now recall that $\frac{\alpha^\tau}{\alpha^{\tau'}} = \exp\left\{Z(\mathbf{m}^\tau) - Z(\mathbf{m}^{\tau'})\right\}$, where $\mathbf{m}^\tau$ and $\mathbf{m}^{\tau'}$ are the centre points of τ, resp. τ'. Then (for $\tau \neq \tau'$) $\left(Z(\mathbf{m}^\tau) - Z(\mathbf{m}^{\tau'})\right)$ is a normal random variable with mean 0 and variance

$$\begin{aligned}
\widetilde{\sigma}^2 &= \mathbb{E}\left[\left(Z(\mathbf{m}^\tau) - Z(\mathbf{m}^{\tau'})\right)^2\right] \\
&= \mathbb{E}\left[Z(\mathbf{m}^\tau)^2\right] + \mathbb{E}\left[Z(\mathbf{m}^{\tau'})^2\right] - 2\mathbb{E}\left[Z(\mathbf{m}^\tau) Z(\mathbf{m}^{\tau'})\right] \\
&= 2\sigma^2\left(1 - \exp\left\{-\left\|\mathbf{m}^\tau - \mathbf{m}^{\tau'}\right\|_2 / \lambda\right\}\right)
\end{aligned} \tag{5.88}$$

for fields with covariance of Ornstein-Uhlenbeck type.

As $\left\|\mathbf{m}^\tau - \mathbf{m}^{\tau'}\right\|_2 \leq H$ for two points $\mathbf{m}^\tau, \mathbf{m}^{\tau'} \in \Omega_j$, we obtain

$$\begin{aligned}
\widetilde{\sigma}^2 &= 2\sigma^2\left(1 - \exp\left\{-\left\|\mathbf{m}^\tau - \mathbf{m}^{\tau'}\right\|_2 / \lambda\right\}\right) \\
&\leq 2\sigma^2\left\{1 - \exp\left\{-H/\lambda\right\}\right\}.
\end{aligned} \tag{5.89}$$

Combining these bounds, we get

$$\begin{aligned}\mathbb{E}\left[\max_j \max_{\tau,\tau'\in\mathcal{T}^h(\bar{\Omega}_j)}\left(\frac{\alpha^\tau}{\alpha^{\tau'}}\right)\right] &\leq \sum_j \sum_{\tau\in\mathcal{T}^h(\bar{\Omega}_j)} \sum_{\tau'\in\mathcal{T}^h(\bar{\Omega}_j)} \mathbb{E}\left[\frac{\alpha^\tau}{\alpha^{\tau'}}\right] \\ &\leq \sum_j \sum_{\tau\in\mathcal{T}^h(\bar{\Omega}_j)} \sum_{\tau'\in\mathcal{T}^h(\bar{\Omega}_j)} \exp\left\{1/2\cdot\widetilde{\sigma}^2\right\} \qquad \text{by (3.7)} \\ &\leq \sum_j \sum_{\tau\in\mathcal{T}^h(\bar{\Omega}_j)} \sum_{\tau'\in\mathcal{T}^h(\bar{\Omega}_j)} \exp\left\{\sigma^2\left\{1-\exp\left\{-H/\lambda\right\}\right\}\right\} \\ &= p\cdot\left(\#\mathcal{T}^h_{\Omega_j}\right)^2\cdot\exp\left\{\sigma^2\left\{1-\exp\left\{-H/\lambda\right\}\right\}\right\}.\end{aligned} \tag{5.90}$$

□

In a very similar way, if $\#\mathcal{T}^h_{\omega_{\Omega_j}}$ is the number of fine grid elements in $\bar{\omega}_{\Omega_j}$, we get

LEMMA 5.19
With $3H$ being an upper bound for the maximum diameter of the neighbourhoods of subdomains ω_{Ω_j}, $j=1,...,p$, we obtain

$$\mathbb{E}\left[\max_j \max_{\tau,\tau'\in\mathcal{T}^h(\bar{\omega}_{\Omega_j})}\left(\frac{\alpha^\tau}{\alpha^{\tau'}}\right)\right] \leq p\cdot\left(\#\mathcal{T}^h_{\omega_{\Omega_j}}\right)^2\cdot\exp\left\{\sigma^2\left\{1-\exp\left\{-3H/\lambda\right\}\right\}\right\}. \tag{5.91}$$

Combining these two lemmata with the results from the previous subsection, we obtain

THEOREM 5.20
The expectation of the condition number $\kappa\left(M^{-1}_{2,Lin}A(\alpha)\right)$ for fields α of Ornstein-Uhlenbeck type can be bounded for a constant C (independent of α, h and H) and with the notations as before by

$$\mathbb{E}\left[\kappa\left(M^{-1}_{2,Lin}A(\alpha)\right)\right] \leq C\left(1+\beta^{-1}\right)p\left(\#\mathcal{T}^h_{\Omega_j}\right)^2\exp\left\{\sigma^2\cdot\left\{1-\exp\left\{-H/\lambda\right\}\right\}\right\}B(d), \tag{5.92}$$

resp. by

$$\mathbb{E}\left[\kappa\left(M^{-1}_{2,Lin}A(\alpha)\right)\right] \leq C\left(1+\beta^{-1}\right)p\left(\#\mathcal{T}^h_{\omega_{\Omega_j}}\right)^2\exp\left\{\sigma^2\left\{1-\exp\left\{-3H/\lambda\right\}\right\}\right\}. \tag{5.93}$$

5.3.2 Probability estimates based on Poisson clumping heuristics

After these bounds for general random fields, we are now interested in probability distributions for the condition numbers with underlying field of Ornstein-Uhlenbeck type with correlation length λ in the 1D and 2D case.
As in Chapter 3 and Chapter 4 denote

$$f_0(b) := (2\pi)^{-d/2}\sigma^{-1}\exp\left\{-\frac{1}{2}db^2/\sigma^2\right\}, \qquad b \in \mathbb{R}. \tag{5.94}$$

Then as shown in Chapter 4 the probability that the maximum of the random field in a subregion Ω_j is less than a fixed value b is for one-dimensional fields (i.e. $d = 1$) given by

$$\mathbb{P}\left[\sup_{\mathbf{x}\in\Omega_j}(\alpha(\mathbf{x})) \leq \exp\{b\}\right] \approx \exp\left\{-f_0(b)\cdot\lambda^{-1}b\,|\Omega_j|\right\}, \tag{5.95}$$

and for two-dimensional fields (i.e. $d = 2$) bounded by

$$\begin{aligned}\exp\left\{-f_0(b)\cdot\lambda^{-2}b^3\,|\Omega_j|\right\} &\lesssim \mathbb{P}\left[\sup_{\mathbf{x}\in\Omega_j}(\alpha(\mathbf{x})) \leq \exp\{b\}\right] \\ &\lesssim \exp\left\{-f_0(b)\cdot\lambda^{-2}b^3\frac{|\Omega_j|}{2}\right\}.\end{aligned} \tag{5.96}$$

Furthermore denote (as in Chapter 3)

$$f_1(b) := 2^{-1}(2\pi)^{-d/2}\sigma^{-1}\exp\left\{-\frac{1}{8}db^2/\sigma^2\right\}, \qquad b \in \mathbb{R}. \tag{5.97}$$

Equivalently to (4.28) we can then show for $d = 1$

$$\begin{aligned}&\exp\left\{-2f_1(b)\cdot\lambda^{-1}b|\Omega_j|\right\} \\ &\lesssim \mathbb{P}\left[\frac{\sup_{\mathbf{x}\in\Omega_j}(\alpha(\mathbf{x}))}{\inf_{\mathbf{y}\in\Omega_j}(\alpha(\mathbf{y}))} \leq \exp\{b\}\right] \\ &\lesssim 2\exp\left\{-f_1(b)\cdot\lambda^{-1}b|\Omega_j|\right\} - \exp\left\{-2f_1(b)\cdot\lambda^{-1}b|\Omega_j|\right\}\end{aligned} \tag{5.98}$$

and equivalently to (4.29) for $d = 2$

$$\begin{aligned}&\exp\left\{-2f_1(b)\cdot\lambda^{-2}b^3|\Omega_j|\right\} \\ &\lesssim \mathbb{P}\left[\frac{\sup_{\mathbf{x}\in\Omega_j}(\alpha(\mathbf{x}))}{\inf_{\mathbf{y}\in\Omega_j}(\alpha(\mathbf{y}))} \leq \exp\{b\}\right] \\ &\lesssim 2\exp\left\{-f_1(b)\cdot\lambda^{-2}b^3\frac{|\Omega_j|}{2}\right\} - \exp\left\{-2f_1(b)\cdot\lambda^{-2}b^3|\Omega_j|\right\}.\end{aligned} \tag{5.99}$$

We now know:

$$\begin{aligned}\prod_{j=1}^{p}\mathbb{P}\left[\left(\frac{\sup_{\mathbf{x}\in\Omega_j}(\alpha(\mathbf{x}))}{\inf_{\mathbf{x}\in\Omega_j}(\alpha(\mathbf{x}))}\right)\leq\exp\{b\}\right]&\lesssim\mathbb{P}\left[\sup_j\frac{\sup_{\mathbf{x}\in\Omega_j}(\alpha(\mathbf{x}))}{\inf_{\mathbf{x}\in\Omega_j}(\alpha(\mathbf{x}))}\leq\exp\{b\}\right]\\&\lesssim\mathbb{P}\left[\left(\frac{\sup_{\mathbf{x}\in\Omega_k}(\alpha(\mathbf{x}))}{\inf_{\mathbf{x}\in\Omega_j}(\alpha(\mathbf{x}))}\right)\leq\exp\{b\}\right],\end{aligned}\tag{5.100}$$

for any subdomain Ω_k.
Let us now consider the following case:

Let us assume that the correlation length λ is small compared to the subdomain size H. In this case we can then assume that the maximum ratios in the different subdomains are independent of each other. We then know

$$\mathbb{P}\left[\sup_j\frac{\sup_{\mathbf{x}\in\Omega_j}(\alpha(\mathbf{x}))}{\inf_{\mathbf{x}\in\Omega_j}(\alpha(\mathbf{x}))}\leq\exp\{b\}\right]\approx\prod_{j=1}^{p}\mathbb{P}\left[\left(\frac{\sup_{\mathbf{x}\in\Omega_j}(\alpha(\mathbf{x}))}{\inf_{\mathbf{x}\in\Omega_j}(\alpha(\mathbf{x}))}\right)\leq\exp\{b\}\right].\tag{5.101}$$

Combining (5.101) and (5.98) we therefore obtain for $d=1$

$$\begin{aligned}&\prod_{j=1}^{p}\left(\exp\left\{-2f_1(b)\cdot\lambda^{-1}b|\Omega_j|\right\}\right)\\&\lesssim\mathbb{P}\left[\sup_j\frac{\sup_{\mathbf{x}\in\Omega_j}(\alpha(\mathbf{x}))}{\inf_{\mathbf{y}\in\Omega_j}(\alpha(\mathbf{y}))}\leq\exp\{b\}\right]\\&\lesssim\prod_{j=1}^{p}\left(2\exp\left\{-f_1(b)\cdot\lambda^{-1}b|\Omega_j|\right\}-\exp\left\{-2f_1(b)\cdot\lambda^{-1}b|\Omega_j|\right\}\right).\end{aligned}\tag{5.102}$$

Similarly combining (5.101) and (5.99) we get for $d=2$

$$\begin{aligned}&\prod_{j=1}^{p}\left(\exp\left\{-2f_1(b)\cdot\lambda^{-2}b^3|\Omega_j|\right\}\right)\\&\lesssim\mathbb{P}\left[\sup_j\frac{\sup_{\mathbf{x}\in\Omega_j}(\alpha(\mathbf{x}))}{\inf_{\mathbf{y}\in\Omega_j}(\alpha(\mathbf{y}))}\leq\exp\{b\}\right]\\&\lesssim\prod_{j=1}^{p}\left(2\exp\left\{-f_1(b)\cdot\lambda^{-2}b^3\frac{|\Omega_j|}{2}\right\}-\exp\left\{-2f_1(b)\cdot\lambda^{-2}b^3|\Omega_j|\right\}\right).\end{aligned}\tag{5.103}$$

Equivalently we can show, if we assume that the maximum ratios in the different

neighbourhoods ω_{Ω_j} of subdomains Ω_j, $j = 1, ..., p$, are independent of each other (which can be a pretty strong assumption, since these neighbourhoods overlap), for $d = 1$

$$\begin{aligned}
&\prod_{j=1}^{p} \left(\exp\left\{-2f_1(b) \cdot \lambda^{-1} b|\omega_{\Omega_j}|\right\}\right) \\
&\lesssim \mathbb{P}\left[\sup_j \frac{\sup_{\mathbf{x}\in\omega_{\Omega_j}}(\alpha(\mathbf{x}))}{\inf_{\mathbf{y}\in\omega_{\Omega_j}}(\alpha(\mathbf{y}))} \leq \exp\{b\}\right] \\
&\lesssim \prod_{j=1}^{p} \left(2\exp\left\{-f_1(b) \cdot \lambda^{-1} b|\omega_{\Omega_j}|\right\} - \exp\left\{-2f_1(b) \cdot \lambda^{-1} b|\omega_{\Omega_j}|\right\}\right)
\end{aligned} \tag{5.104}$$

and for $d = 2$

$$\begin{aligned}
&\prod_{j=1}^{p} \left(\exp\left\{-2f_1(b) \cdot \lambda^{-2} b^3|\omega_{\Omega_j}|\right\}\right) \\
&\lesssim \mathbb{P}\left[\sup_j \frac{\sup_{\mathbf{x}\in\omega_{\Omega_j}}(\alpha(\mathbf{x}))}{\inf_{\mathbf{y}\in\omega_{\Omega_j}}(\alpha(\mathbf{y}))} \leq \exp\{b\}\right] \\
&\lesssim \prod_{j=1}^{p} \left(2\exp\left\{-f_1(b) \cdot \lambda^{-2} b^3 \frac{|\omega_{\Omega_j}|}{2}\right\} - \exp\left\{-2f_1(b) \cdot \lambda^{-2} b^3|\omega_{\Omega_j}|\right\}\right).
\end{aligned} \tag{5.105}$$

Using bound (5.102), resp. (5.103) in Theorem 5.16 and bound (5.104), resp. (5.105) in Theorem 5.17 for small correlation length (compared to the subdomain size) in a similar way as in inequality (4.30) then leads to bounds of *probability distributions for the condition numbers of the domain decomposition preconditioned systems* with underlying permeability fields of Ornstein-Uhlenbeck type. This makes it possible to predict, in how many cases the condition numbers (and therefore the conjugate gradient iteration numbers) will probably stay below a certain value, which makes the computation costs, which are involved, much more predictable. These theoretical results can therefore be quite significant in the practical use of domain decomposition preconditioners for the groundwater flow problem.

5.4 Numerical results

We now would like to check numerically how good the bounds for the probability distributions of the local maxima $\sup_j \frac{\sup_{\mathbf{x}\in\Omega_j}(\alpha(\mathbf{x}))}{\inf_{\mathbf{x}\in\Omega_j}(\alpha(\mathbf{x}))}$ based on Poisson clumping heuristics (see (5.102) for $d = 1$ and (5.103) for $d = 2$) are in practice. Let us as an example consider the 2D case for $\sigma^2 = 4$, $h = 1/256$ and $H = 1/4$ and various correlation lengths λ. We then obtain Table 5.1.

λ	**Prob.** $\hat{\mathbb{P}}[\mathbf{b}]$	**Theoretical lower bound** $\exp\{\mathbf{b}\}$	**Numerical value** $\exp\left\{\tilde{\mathbf{b}}\right\}$	**Theoretical upper bound** $\exp\{\mathbf{B}\}$
1/32	85%	$2.19 \cdot 10^4$	$2.45 \cdot 10^6$	$2.55 \cdot 10^6$
	87.5%	$2.20 \cdot 10^4$	$2.78 \cdot 10^6$	$2.87 \cdot 10^6$
	90%	$2.20 \cdot 10^4$	$3.17 \cdot 10^6$	$3.31 \cdot 10^6$
	92.5%	$2.21 \cdot 10^4$	$3.82 \cdot 10^6$	$3.95 \cdot 10^6$
	95%	$2.21 \cdot 10^4$	$4.97 \cdot 10^6$	$5.04 \cdot 10^6$
	97.5%	$2.22 \cdot 10^4$	$7.41 \cdot 10^6$	$7.53 \cdot 10^6$
	99%	$2.22 \cdot 10^4$	$7.58 \cdot 10^6$	$1.25 \cdot 10^7$
1/64	85%	$8.22 \cdot 10^4$	$5.41 \cdot 10^6$	$5.76 \cdot 10^6$
	87.5%	$8.23 \cdot 10^4$	$6.12 \cdot 10^6$	$6.44 \cdot 10^6$
	90%	$8.25 \cdot 10^4$	$6.46 \cdot 10^6$	$7.36 \cdot 10^6$
	92.5%	$8.26 \cdot 10^4$	$8.30 \cdot 10^6$	$8.71 \cdot 10^6$
	95%	$8.28 \cdot 10^4$	$8.99 \cdot 10^6$	$1.09 \cdot 10^7$
	97.5%	$8.29 \cdot 10^4$	$1.08 \cdot 10^7$	$1.60 \cdot 10^7$
	99%	$8.30 \cdot 10^4$	$1.42 \cdot 10^7$	$2.58 \cdot 10^7$
1/128	85%	$2.53 \cdot 10^5$	$5.77 \cdot 10^6$	$1.24 \cdot 10^7$
	87.5%	$2.54 \cdot 10^5$	$6.95 \cdot 10^6$	$1.38 \cdot 10^7$
	90%	$2.54 \cdot 10^5$	$7.88 \cdot 10^6$	$1.57 \cdot 10^7$
	92.5%	$2.54 \cdot 10^5$	$8.70 \cdot 10^6$	$1.83 \cdot 10^7$
	95%	$2.55 \cdot 10^5$	$9.53 \cdot 10^6$	$2.28 \cdot 10^7$
	97.5%	$2.55 \cdot 10^5$	$1.51 \cdot 10^7$	$3.26 \cdot 10^7$
	99%	$2.56 \cdot 10^5$	$1.88 \cdot 10^7$	$5.16 \cdot 10^7$
1/256	85%	$6.88 \cdot 10^5$	$6.54 \cdot 10^6$	$2.57 \cdot 10^7$
	87.5%	$6.89 \cdot 10^5$	$7.49 \cdot 10^6$	$2.84 \cdot 10^7$
	90%	$6.90 \cdot 10^5$	$8.26 \cdot 10^6$	$3.20 \cdot 10^7$
	92.5%	$6.91 \cdot 10^5$	$9.46 \cdot 10^6$	$3.72 \cdot 10^7$
	95%	$6.92 \cdot 10^5$	$1.02 \cdot 10^7$	$4.58 \cdot 10^7$
	97.5%	$6.93 \cdot 10^5$	$1.74 \cdot 10^7$	$6.45 \cdot 10^7$
	99%	$6.94 \cdot 10^5$	$2.84 \cdot 10^7$	$1.00 \cdot 10^8$

Table 5.1: Theoretical and computed probability distributions for $\sigma^2 = 4$, $h = 1/256$ and $H = 1/4$

For $\sigma^2 = 4$, $h = 1/128$ and $H = 1/8$ (i.e. larger subdomains) we obtain Table 5.2:

λ	**Prob.** $\hat{\mathbb{P}}$[b]	**Theoretical lower bound** $\exp\{\mathbf{b}\}$	**Numerical value** $\exp\left\{\tilde{\mathbf{b}}\right\}$	**Theoretical upper bound** $\exp\{\mathbf{B}\}$
1/16	85%	$1.17 \cdot 10^5$	$1.85 \cdot 10^5$	$1.06 \cdot 10^6$
	87.5%	$1.22 \cdot 10^5$	$2.26 \cdot 10^5$	$1.20 \cdot 10^6$
	90%	$1.26 \cdot 10^5$	$2.75 \cdot 10^5$	$1.40 \cdot 10^6$
	92.5%	$1.31 \cdot 10^5$	$3.58 \cdot 10^5$	$1.70 \cdot 10^6$
	95%	$1.36 \cdot 10^5$	$5.38 \cdot 10^5$	$2.21 \cdot 10^6$
	97.5%	$1.42 \cdot 10^5$	$8.97 \cdot 10^5$	$3.39 \cdot 10^6$
	99%	$1.48 \cdot 10^5$	$1.78 \cdot 10^6$	$5.80 \cdot 10^6$
1/8	85%	$3.36 \cdot 10^4$	$3.67 \cdot 10^4$	$4.07 \cdot 10^5$
	87.5%	$3.51 \cdot 10^4$	$3.99 \cdot 10^4$	$4.69 \cdot 10^5$
	90%	$3.66 \cdot 10^4$	$4.40 \cdot 10^4$	$5.54 \cdot 10^5$
	92.5%	$3.83 \cdot 10^4$	$5.80 \cdot 10^4$	$6.83 \cdot 10^5$
	95%	$4.01 \cdot 10^4$	$8.26 \cdot 10^4$	$9.07 \cdot 10^5$
	97.5%	$4.21 \cdot 10^4$	$1.43 \cdot 10^5$	$1.44 \cdot 10^6$
	99%	$4.34 \cdot 10^4$	$2.86 \cdot 10^5$	$2.56 \cdot 10^6$
1/4	85%	$7.14 \cdot 10^3$	$7.41 \cdot 10^3$	$1.41 \cdot 10^5$
	87.5%	$7.54 \cdot 10^3$	$7.99 \cdot 10^3$	$1.65 \cdot 10^5$
	90%	$7.98 \cdot 10^3$	$8.93 \cdot 10^3$	$1.99 \cdot 10^5$
	92.5%	$8.46 \cdot 10^3$	$9.84 \cdot 10^3$	$2.51 \cdot 10^5$
	95%	$8.98 \cdot 10^3$	$1.09 \cdot 10^4$	$3.43 \cdot 10^5$
	97.5%	$9.56 \cdot 10^3$	$1.78 \cdot 10^4$	$5.70 \cdot 10^5$
	99%	$9.93 \cdot 10^3$	$3.97 \cdot 10^4$	$1.07 \cdot 10^6$
1/2	85%	$5.92 \cdot 10^2$	$6.64 \cdot 10^2$	$4.17 \cdot 10^4$
	87.5%	$6.71 \cdot 10^2$	$7.83 \cdot 10^2$	$5.01 \cdot 10^4$
	90%	$7.64 \cdot 10^2$	$9.10 \cdot 10^2$	$6.23 \cdot 10^4$
	92.5%	$8.44 \cdot 10^2$	$1.11 \cdot 10^3$	$8.13 \cdot 10^4$
	95%	$9.32 \cdot 10^2$	$1.50 \cdot 10^3$	$1.16 \cdot 10^5$
	97.5%	$1.24 \cdot 10^3$	$2.20 \cdot 10^3$	$2.05 \cdot 10^5$
	99%	$1.73 \cdot 10^3$	$5.70 \cdot 10^3$	$4.10 \cdot 10^5$

Table 5.2: Theoretical and computed probability distributions for $\sigma^2 = 4$, $h = 1/128$ and $H = 1/8$

These results show that the bounds are not as close as those in Tables 3.5 and 4.2, but that they still give a good idea of the distributions of $\mathbb{P}\left[\sup_j \frac{\sup_{\mathbf{x}\in\Omega_j}(\alpha(\mathbf{x}))}{\inf_{\mathbf{y}\in\Omega_j}(\alpha(\mathbf{y}))} \leq \exp\{b\}\right]$ for large levels b.

We next would like to study the dependence of the domain decomposition preconditioners on the parameters h, H, β, σ and λ for two dimensional domains in practice for problems with underlying lognormal random fields of Ornstein-Uhlenbeck type. The random fields used for simulations in practice normally have a variance ranging between $\sigma^2 = 1$ to $\sigma^2 = 16$. Our numerical experiments are therefore performed for variances of this order.

Let us first focus on the dependence on the subdomain size H. All results presented here are averages over 100 random fields. The random fields were computed using the Gaussian library by Boris Kozintsev [84, 85]. We obtained for $h = 1/128$ and $\lambda = 10h$ for overlap h, resp. $2h$ for additive Schwarz preconditioners without coarse grids (1 Level) and with linear coarsening (2 Level) the results summarized in Tables 5.3 and 5.4.

Var. σ^2	**H = 1/32 :**		**H = 1/16 :**		**H = 1/8 :**	
	1 Level	**2 Level**	**1 Level**	**2 Level**	**1 Level**	**2 Level**
0	116	20	89	27	77	35
1	241	27	197	36	143	49
2	317	31	241	42	171	58
4	396	37	316	52	219	73
6	496	44	369	63	259	89
8	589	52	444	71	298	105
16	920	81	695	125	454	185
24	$> 1,000$	124	930	210	634	294

Table 5.3: CG iteration numbers - Fields of Ornstein-Uhlenbeck type with $\lambda = 10h$ ($h = 1/128$ and overlap$= h$)

Var. σ^2	**H = 1/32 :**		**H = 1/16 :**		**H = 1/8 :**	
	1 Level	**2 Level**	**1 Level**	**2 Level**	**1 Level**	**2 Level**
0	78	17	69	22	55	29
1	153	21	140	30	110	38
2	183	24	161	34	128	45
4	230	28	198	42	160	57
6	273	33	244	49	187	67
8	311	36	288	60	210	82
16	438	56	358	101	296	144
24	546	88	479	163	369	219

Table 5.4: CG iteration numbers - Fields of Ornstein-Uhlenbeck type with $\lambda = 10h$ ($h = 1/128$ and overlap$= 2h$)

We can first of all see, that the two-level additive Schwarz preconditioner preforms better than the one-level preconditioner in terms of conjugate gradient iteration numbers in all cases. The setup time and the time per iteration for the two-level method are slightly larger than for the one-level preconditioner. This however does not change the fact, that for all experiments considered in this thesis, the two-level additive Schwarz preconditioner was always superior in terms of overall computation times to the one-level preconditioner.
The second thing we can see from these experiments is, that a reduction of the subdomain size H increases the number of cg-iterations when no coarse grid is used, but decreases the number of cg-iterations for the two-level additive Schwarz preconditioner with linear coarsening. An increase of the the overlap reduces the number of iterations here.

Focusing on the two-level additive Schwarz preconditioner we can therefore clearly reduce the number of iterations by reducing the subdomain size and increasing the overlap (as the theoretical considerations also suggested). However these changes come with an increase of computation time both for the setup of the preconditioner and per iteration.
Furthermore we observe that the number of iterations grows strongly with the variance σ^2. We will try to reduce this dependence by the introduction of a new coarsening operator in Chapter 6 of this thesis.
An exponential dependence on σ^2 can also be observed for fields with no correlation, as Tables 5.5 and 5.6 show.

Var. σ^2	**H = 1/32 :**		**H = 1/16 :**		**H = 1/8 :**	
	1 Level	**2 Level**	**1 Level**	**2 Level**	**1 Level**	**2 Level**
0	116	20	89	27	77	35
1	256	38	205	48	139	48
2	334	55	270	66	176	65
4	536	91	438	112	279	107
6	867	146	727	167	448	159
8	1,337	238	1,138	245	664	249
16	> 1,000	842	> 1,000	829	> 1,000	938
24	> 1,000	> 1,000	> 1,000	> 1,000	> 1,000	> 1,000

Table 5.5: CG iteration numbers - No covariance ($h = 1/128$ and overlap= h)

Next we would like to study the dependence of the conjugate gradient iteration numbers on the correlation length of the underlying fields in more detail. We therefore fix

Var. σ^2	H = 1/32 :		H = 1/16 :		H = 1/8 :	
	1 Level	2 Level	1 Level	2 Level	1 Level	2 Level
0	78	17	69	22	55	29
1	136	23	126	27	95	32
2	156	30	130	34	107	37
4	183	43	161	54	122	55
6	212	59	183	72	140	77
8	245	80	210	96	154	101
16	446	211	391	272	306	266
24	789	444	729	608	486	467

Table 5.6: CG iteration numbers - No covariance ($h = 1/128$ and overlap= $2h$)

$h = 1/128$ and $H = 1/16$ and vary λ. The results can be found in Tables 5.7 and 5.8.

Var. σ^2	$\lambda = 1/2$:		$\lambda = 1/4$:		$\lambda = 1/8$:	
	1 Level	2 Level	1 Level	2 Level	1 Level	2 Level
1	116	25	120	26	134	28
2	122	26	132	28	158	30
4	131	28	151	31	201	37
6	139	30	170	35	250	43
8	148	32	186	38	256	50
12	162	35	212	45	300	66
16	178	39	241	50	342	86

Table 5.7: CG iteration - Dependence on the correlation length λ ($h = 1/128$, $H = 1/16$, Overlap $\beta H = 2h$)

Var. σ^2	$\lambda = 1/16$:		$\lambda = 1/32$:		$\lambda = 1/64$:	
	1 Level	2 Level	1 Level	2 Level	1 Level	2 Level
1	146	29	150	31	153	32
2	169	35	175	40	189	40
4	229	47	247	59	263	63
6	296	62	309	85	367	96
8	336	79	392	118	474	130
12	424	121	628	210	743	245
16	578	178	846	340	$> 1,000$	407

Table 5.8: CG iteration - Dependence on the correlation length λ ($h = 1/128$, $H = 1/16$, Overlap $\beta H = 2h$)

We can see that the number of iterations increases strongly for decreasing λ especially for large variance. This effect can also be reduced by using a coarsening operator that depends on the underlying field as we will show in the next chapter.

5.5 Summary

In this chapter we did some first steps in finding a preconditioner for discretisations of the second order partial differential equation (5.1, 5.2). We first of all introduced the Schwarz alternating method for the case of two overlapping subdomains. Next we moved on to a two level additive Schwarz method by introducing a coarse grid in such a way, that each fine grid element lies in exactly one coarse grid element $\bar{\Omega}_j$, $j = 1, ..., p$. A space $\mathcal{V}_j$ for every overlapping subdomain $\widetilde{\Omega_j}$ and one, $\mathcal{V}_0$, on the coarse grid were then introduced. We obtained the one level additive Schwarz preconditioner $M_1^{-1} := \sum_{j=1}^{p} R_j^T A_j^{-1} R_j$ and the two level additive Schwarz preconditioner $M_{2,Lin}^{-1} := \sum_{j=0}^{p} R_j^T A_j^{-1} R_j$. As the matrices A_j, $j = 0, ...p$, are of significantly smaller size than the original stiffness matrix, $A(\boldsymbol{\alpha})$, applying these preconditioners can potentially lead to reduced computational costs in iterative methods, if M_1^{-1}, resp. $M_{2,Lin}^{-1}$ approximate $A(\boldsymbol{\alpha})^{-1}$ well enough.

We then focused on the convergence theory of additive Schwarz preconditioners by first quoting a bound for the condition number of the stiffness matrix preconditioned by the one-level method (Theorem 5.5). Using a colouring argument we next showed for the two-level preconditioner with linear coarsening that the maximum eigenvalue of the preconditioned system can be bounded by a constant. The minimum eigenvalue however can depend strongly on the underlying permeability field and a general bound for it was given using the partition property of the subspaces $\mathcal{V}_j$. In the study of domain decomposition preconditioners not much research had been done on this dependence so far. The aim of the following discussion therefore was to find out more about this dependence and to show that it is weaker than the linear dependence on $\left(\max_{\mathbf{x}\in\Omega} \alpha(\mathbf{x}) / \min_{\mathbf{y}\in\Omega} \alpha(\mathbf{y})\right)$ for random fields with large correlation length relative to the size of the subdomains.

We first of all proved that the condition number of the two-level additive Schwarz preconditioned matrix can in the worst case depend linearly on $\max_j \left(\max_{\mathbf{x}\in\Omega_j} \alpha(\mathbf{x}) / \min_{\mathbf{y}\in\Omega_j} \alpha(\mathbf{y})\right)$ as well for two dimensional fields a factor $(1 + \log(H/h))$ and for three dimensional fields a factor H/h. This means that the condition number is independent of jumps of the permeability across subdomain boundaries, as it was shown before among others by Chan and Mathew [28]. However the dependence on the coarse grid size is

strong for three-dimensional problems. We therefore derived a new bound which gets rid of the factor $(1 + \log(H/h))$ for two-dimensional and of the factor H/h for three-dimensional domains, but therefore depends no longer on the ratio of permeabilities in one subdomain Ω_j but on a (sometimes significantly larger) ratio of permeabilities in a neighbourhood ω_{Ω_j} of subdomains.
Next we tackled the dependence on the overlap, βH, of our new bounds by improving the dependence from $\left(1 + \beta^{-2}\right)$ to $\left(1 + \beta^{-1}\right)$. The bounds were finally tested by a series of numerical experiments for a two media case as well as for random fields.
Although, as our new bounds prove, two-level additive Schwarz preconditioners with linear coarsening behave well for fields with medium and large correlation length, the theoretical bounds and numerical results show that the condition numbers of the preconditioned matrices can in the worst case still depend exponentially on the variance of the underlying field.

Chapter 6

Domain decomposition with multiscale interpolation

"No problem can be solved from the same level of consciousness that created it."

— Albert Einstein (German born physicist who developed the special and general theories of relativity. Nobel Prize for Physics in 1921. 1879-1955) —

"If you don't scale the mountain, you can't view the plain."

— Chinese Proverb —

6.1 Multiscale finite elements

6.1.1 Introduction

We have seen in the previous chapter that the convergence of the preconditioned additive Schwarz method with linear coarsening depends strongly on the underlying permeability field. In fact the condition number of the preconditioned system can in the worst case increase linearly with the maximum ratio of the coefficients within the subdomains as shown in Theorem 5.16. One way to reduce this effect for fields with large correlation length (compared with the subdomain size H) therefore might be to reduce the size of the subdomains. This however means that we have to solve bigger coarse

grid problems and that we have to deal with extra storage requirements. A better way therefore might be to find a coarse grid operator which takes into consideration the fine scale information of the problem.

Various numerical methods were developed over the last years capturing these small scale properties. Some of the most relevant ones are:

- **Wavelet homogenization techniques:** The aim here is to eliminate the small scales. To achieve this the numerical solution is represented using a wavelet basis (see therefore Dorobantu and Engquist [39]). The resulting operators become dense, but can be approximated well in sparse form, as shown by Engquist and Runborg [55].

- **Heterogeneous multiscale methods:** These methods try to couple macroscopic and microscopic models effectively (as described for example by E [48], E and Engquist [49, 50, 51]). If the macroscopic model is not known explicitly, the microscopic solver is used to supply the necessary data. By exploiting the scale separation, the complexity of the microscopic solver can be reduced significantly.

- **Residual-Free Bubble methods:** This method is closely related to the multiscale finite element method, used in this thesis, as it also enriches the Galerkin finite element space with (bubble) functions, that have local support on every element of a given triangulation. The bubbles are then eliminated step by step leaving a generalised Galerkin scheme, which often shows improved approximation properties. The resulting scheme is locally residual free, i.e. it is a solution of the considered partial differential equation inside each element. Therefore only properties inside each element are resolved and additional functions have to be added to the bubble space to be able to resolve features across boundaries. Details about this method are discussed by Hughes [78], Sangalli [115] and Cangiani and Suli [25].

- **Multiscale Finite Element Methods (MsFEM):** This method is an extension of the Generalized Finite Element Method, that was first discussed by Babuška and Osborn [6] and by Babuška et al. [7]. Its aim is to incorporate the fine scale information into coarse grid finite element basis functions. By the reduction of the number of basis functions, computation costs can be reduced significantly. Details of the MsFE method will be given in this section.

The multiscale finite element method itself was developed by Hou and Wu [75] for problems arising from composite materials and flows in porous media. It was used to

improve the exactness of the solution, while keeping the size of the system, that has to be solved, small. Parallelisation strategies were given and a first approach to its analysis for periodic problems was presented.
The behaviour of the method for problems with strongly varying coefficients was investigated by Hou et al. [76]. The conforming method often shows some resonance errors, which might cause problems with respect to the accuracy in some cases. Efendiev et al. [54] therefore considered a nonconforming multiscale finite element method, also known as *oversampling.*
The multiscale finite element method with and without oversampling and a detailed analysis of it was also studied in Yalchin Efendiev's Ph.D. thesis (see Efendiev [52]) and various 2D and 3D numerical results are given.
Efendiev et al. [53] showed how the method can be used in practice for the modeling of transport problems in heterogeneous media.
Motivated by the simulation of flow transport through heterogeneous media, Chen and Hou [29] presented a mixed finite element version of the MsFEM, using an oversampling technique and solving Neumann boundary problems for the construction of the basis functions.
Good summaries of the method with some new error bounds and approaches for its analysis can also be found in Hou [73, 74].

6.1.2 Construction of multiscale basis functions

As before let us consider the second order partial differential equation

$$-\mathrm{div}\left(\alpha\left(\mathbf{x}\right)\boldsymbol{\nabla} u\left(\mathbf{x}\right)\right) = f(\mathbf{x}) \text{ for } \mathbf{x} \in \Omega \subset \mathbb{R}^d, \tag{6.1}$$

assuming for simplicity $u(\mathbf{x}) = 0$ on Γ. This problem will be considered for general dimension $d \geq 1$ first of all without going into any detail and then more precisely separately for $d = 1$ and for $d > 1$. Let us again introduce a fine grid $\mathcal{T}^h(\bar{\Omega})$ of elements of maximum diameter h which is assumed to be a refinement of a coarse grid $\mathcal{T}^H(\bar{\Omega})$ of elements of maximum diameter H. Thus every element $\tau \in \mathcal{T}^h(\bar{\Omega}_j)$ belongs to exactly one coarse grid element $\bar{\Omega}_j \in \mathcal{T}^H(\bar{\Omega})$. We next define a finite-dimensional subspace of multiscale finite elements on $H_0^1(\Omega)$, the degrees of freedom for which are the nodes of the coarse grid $\mathcal{T}^H(\bar{\Omega})$.
Now denote $\mathcal{N}^H(\Omega)$ the coarse grid degrees of freedom on Ω. The *multiscale coarse grid basis functions* ψ_i^{Ms}, $\mathbf{x}_i^H \in \mathcal{N}^H(\Omega)$, will then be defined on the whole of Ω with

support only on

$$\omega_i := \text{interior}\left(\bigcup_{j:\mathbf{x}_i^H \in \bar{\Omega}_j} \bar{\Omega}_j\right), \tag{6.2}$$

i.e. the union of the subdomains which have $\mathbf{x}_i^H$ as a node (see Figure 6-1 and as defined before in (5.15)). The functions ψ_i^{Ms} will be piecewise linear with respect to the fine mesh and their definition will depend on the underlying permeability field α. For their construction we first of all define *local multiscale basis functions* $\vartheta_{i,\Omega_j}^{Ms}$ on a typical coarse grid element $\bar{\Omega}_j \subset \bar{\omega}_i$.

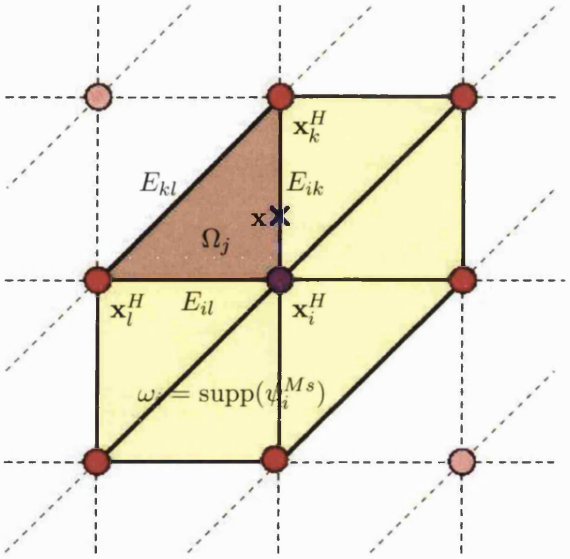


Figure 6-1: Neighbourhood ω_i of a coarse grid node $\mathbf{x}_i^H$ (ω_i is also the support of ψ_i^{Ms}) and notations for one of its subdomains Ω_j (which is also the support of $\vartheta_{i,\Omega_j}^{Ms}$, $\vartheta_{k,\Omega_j}^{Ms}$ and $\vartheta_{l,\Omega_j}^{Ms}$) for $d = 2$

We require that $\vartheta_{i,\Omega_j}^{Ms}$ with $\mathbf{x}_i^H \in \bar{\Omega}_j$ is a finite element solution of the *local problem*

$$-\text{div}\left(\alpha\left(\mathbf{x}\right)\boldsymbol{\nabla} u\left(\mathbf{x}\right)\right) = 0 \quad \text{for } \mathbf{x} \in \Omega_j \tag{6.3}$$

and that it satisfies some suitable boundary conditions on the boundary Γ_j of Ω_j, that we will discuss in more detail later in this section. One possible simple choice for example could be to set $\vartheta_{i,\Omega_j}^{Ms}|_{\Gamma_j} := \psi_i^{Lin}|_{\Gamma_j}$, where $\psi_i^{Lin}|_{\Gamma_j}$ is the restriction of the piecewise linear coarse grid basis function ψ_i^{Lin} to Γ_j, i.e. $\vartheta_{i,\Omega_j}^{Ms}(\mathbf{x}_k^H) := \delta_{ik}$, where $\mathbf{x}_k^H$

is a coarse grid node in $\bar{\Omega}_j$, and linear interpolation between these nodes on the edges of $\bar{\Omega}_j$. As we will see, the quality of the discretisation often depends critically on the boundary conditions.
After having introduced the local basis functions, $\vartheta^{Ms}_{i,\Omega_j}$, for each $\Omega_j \subset \omega_i$, the global multiscale basis function ψ^{Ms}_i on Ω is given by

$$\psi^{Ms}_i|_{\Omega_j} := \vartheta^{Ms}_{i,\Omega_j}, \quad \text{for each } \Omega_j \subset \omega_i \tag{6.4}$$

and on the boundary of each Ω_j by the boundary conditions for the construction of $\vartheta^{Ms}_{i,\Omega_j}$.
Furthermore bearing in mind (5.15), $\psi^{Ms}_i(\mathbf{x}) := 0$ for all $\mathbf{x} \in \Omega \backslash \omega_i$.
Note that whether the basis functions ψ^{Ms}_i are H^1-conforming or not depends on the compatibility of the boundary conditions across the edges of the subdomains.

Originally these global multiscale basis functions were invented to improve the accuracy of a solution by including information about the underlying field α in the finite element basis (see Hou and Wu [75], Hou et al. [76], Efendiev [52], Efendiev et al. [53, 54], Chen and Hou [29], Hou [73, 74]). More recently however there were some promising first attempts of using them as part of preconditioners (see Aarnes and Hou [1]). Jørg Aarnes and Tom Hou applied the multiscale basis functions for the coarsening operator in a hybrid multiplicative-additive Schwarz method. They gave some interesting numerical results for this hybrid preconditioner as well as some first ideas of a good choice of boundary conditions for the construction of the multiscale basis.
We presented some early steps of the analysis of additive Schwarz domain decomposition preconditioners with MsFE-coarsening (see Graham and Lechner [59]).
The MsFE basis will be used in this chapter to replace the linear coarsening operator by a new operator, that includes the information of the underlying permeability field. Different from the preconditioner discussed in Aarnes and Hou [1], our new multiscale additive Schwarz preconditioner can perform the local and global solves simultaneously, which allows almost perfect parallelisation. Furthermore we develop for the first time a theory that gives sharp bounds for a subclass of permeability fields and extensive numerical results for the two-media case and for random permeability fields of Ornstein-Uhlenbeck type.

Recall that the linear extension operator was defined (see (5.18)) to be the $n_h \times n_H$ matrix

$$R^T_{0,Lin} := \left[\psi^{Lin}_1, ..., \psi^{Lin}_{n_H}\right], \tag{6.5}$$

where $\boldsymbol{\psi}_1^{Lin}, ..., \boldsymbol{\psi}_{n_H}^{Lin}$ are the vectors of values of the piecewise linear coarse grid basis functions at the fine grid nodes and n_h and n_H the number of fine, resp. coarse grid nodes in the interior of Ω.
Similarly we define the *multiscale extension operator* to be the $n_h \times n_H$ matrix

$$R_{0,Ms}^T := \left[\boldsymbol{\psi}_1^{Ms}, ..., \boldsymbol{\psi}_{n_H}^{Ms}\right], \tag{6.6}$$

where $\boldsymbol{\psi}_1^{Ms}, ..., \boldsymbol{\psi}_{n_H}^{Ms}$ are the vectors of evaluations of the global multiscale coarse grid basis functions at the fine grid nodes. The multiscale coarse grid stiffness matrix with multiscale coarsening, $A_{0,Ms}$, is then (in analogy to A_0, see (5.19)) defined by

$$A_{0,Ms} := R_{0,Ms} A R_{0,Ms}^T \tag{6.7}$$

and the *two-level additive Schwarz preconditioner* with multiscale coarsening by

$$M_{2,Ms}^{-1} := \sum_{j=1}^{p} R_j^T A_j^{-1} R_j + R_{0,Ms}^T A_{0,Ms}^{-1} R_{0,Ms} = M_1^{-1} + R_{0,Ms}^T A_{0,Ms}^{-1} R_{0,Ms}. \tag{6.8}$$

In the next two subsections we give some idea as to how the multiscale basis functions behave inside elements in which there are large variations in coefficient values.

6.1.2.2 One dimensional case

In the one dimensional case, problem (6.1) reads

$$\begin{aligned} -\left(\alpha\left(x\right)\left(u(x)\right)'\right)' &= f(x) \text{ on } (0,1), \\ u(0) = u(1) &= 0. \end{aligned} \tag{6.9}$$

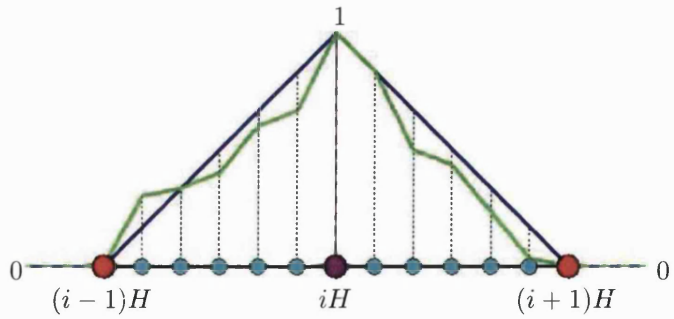


Figure 6-2: Example of a linear basis function ψ_i^{Lin} (blue) and a multiscale basis function ψ_i^{Ms} (green) with support on $[(i-1)H, (i+1)H]$ (fine grid nodes in the support in cyan)

Let us assume that domain $\Omega = (0,1)$ is divided into subdomains $\Omega_i := [(i-1)H, iH]$, $i = 1, ..., p$, $H = 1/p$. Figure 6-2 displays an example for ψ_i^{Ms} and shows that its support consists only of Ω_i and Ω_{i+1}. Assuming Ω_i is divided into a fine mesh the multiscale basis functions $\psi_{i-1}^{Ms}|_{\Omega_i} := \vartheta_{i-1,\Omega_i}^{Ms}$ and $\psi_i^{Ms}|_{\Omega_i} := \vartheta_{i,\Omega_i}^{Ms}$ are finite element solutions with respect to the fine mesh of the PDE

$$\begin{aligned} \left(\alpha\,(x)\,(u(x))'\right)' &= 0 \text{ on } \Omega_i, \\ u((i-1)H) &= 1 \text{ and } u(iH) = 0, \end{aligned} \tag{6.10}$$

resp. of the PDE

$$\begin{aligned} \left(\alpha\,(x)\,(u(x))'\right)' &= 0 \text{ on } \Omega_i, \\ u((i-1)H) &= 0 \text{ and } u(iH) = 1. \end{aligned} \tag{6.11}$$

Now let

$$\mathcal{V}_{0,\Omega_i}^h := \left\{ v : v \text{ is continuous on } \Omega_i, v|_\tau \text{ is linear for } \tau \in \mathcal{T}^h(\bar{\Omega}_i), v((i-1)H) = v(iH) = 0 \right\} \tag{6.12}$$

and

$$\mathcal{V}_{1,\Omega_i}^h := \left\{ v : v \text{ is continuous on } \Omega_i, v|_\tau \text{ is linear for } \tau \in \mathcal{T}^h(\bar{\Omega}_i), v((i-1)H) = 1, v(iH) = 0 \right\}. \tag{6.13}$$

If we then focus on PDE (6.10) for the moment, we search for a solution of the finite element formulation of this problem, i.e. of

$$\text{Find } \vartheta_{i-1,\Omega_i}^{Ms} \in \mathcal{V}_{1,\Omega_i}^h \text{ such that } \int_{\Omega_i} \alpha \left(\vartheta_{i-1,\Omega_i}^{Ms}\right)' v' = 0,\ \forall v \in \mathcal{V}_{0,\Omega_i}^h. \tag{6.14}$$

Let us assume constant coefficients $\alpha = \alpha^\tau$ on each fine grid element $\tau \subset \bar{\Omega}_i$. The piecewise linear solution of this finite element problem on $\Omega_i = [(i-1)H, iH]$ can be found analytically and is defined by the prescription

$$\vartheta_{i-1,\Omega_i}^{Ms}((i-1)H) = 1 \tag{6.15}$$

and for x in the interior of a fine grid element $\tau' \subset \bar{\Omega}_i$

$$\left(\vartheta_{i-1,\Omega_i}^{Ms}(x)\right)' = -\frac{1}{\alpha^{\tau'}} \left(\sum_{\tau \subset \bar{\Omega}_i} \frac{|\tau|}{\alpha^\tau} \right)^{-1}. \tag{6.16}$$

To see that this really describes a solution of problem (6.14), we now only have to prove the following to two points

(i) $\vartheta^{Ms}_{i-1,\Omega_i}(iH) = 0$.
This follows, since

$$\begin{aligned}\vartheta^{Ms}_{i-1,\Omega_i}(iH) &= \vartheta^{Ms}_{i-1,\Omega_i}((i-1)H) + \int_{(i-1)H}^{iH} \left(\vartheta^{Ms}_{i-1,\Omega_i}(x)\right)' \mathrm{d}x \\ &= \vartheta^{Ms}_{i-1,\Omega_i}((i-1)H) + \sum_{\tau' \in \bar{\Omega}_i} -|\tau'| \frac{1}{\alpha^{\tau'}} \left(\sum_{\tau \subset \bar{\Omega}_i} \frac{|\tau|}{\alpha^\tau}\right)^{-1} \\ &= 1 - \left(\sum_{\tau \subset \bar{\Omega}_i} \frac{|\tau|}{\alpha^\tau}\right)\left(\sum_{\tau \subset \bar{\Omega}_i} \frac{|\tau|}{\alpha^\tau}\right)^{-1} = 0.\end{aligned} \tag{6.17}$$

(ii) The piecewise linear function prescribed by (6.15) and (6.16) satisfies the finite element problem. This follows, since $\forall v \in \mathcal{V}^h_{0,\Omega_i}$

$$\begin{aligned}\int_{\Omega_i} \alpha \left(\vartheta^{Ms}_{i-1,\Omega_i}\right)' v' &= \sum_{\tau \subset \bar{\Omega}_i} \int_\tau \alpha^\tau \frac{1}{\alpha^\tau} \left(\sum_{\tau \subset \bar{\Omega}_i} \frac{|\tau|}{\alpha^\tau}\right)^{-1} v' \\ &= \left(\sum_{\tau \subset \bar{\Omega}_i} \frac{|\tau|}{\alpha^\tau}\right)^{-1} (v(iH) - v((i-1)H)) = 0,\end{aligned} \tag{6.18}$$

Similarly one can show that $\vartheta^{Ms}_{i,\Omega_i}$ can be prescribed on Ω_i by

$$\vartheta^{Ms}_{i,\Omega_i}(iH) = 1 \tag{6.19}$$

and for any point x in the interior of a fine grid element $\tau' \subset \bar{\Omega}_i$

$$\left(\vartheta^{Ms}_{i,\Omega_i}(x)\right)' = \frac{1}{\alpha^{\tau'}} \left(\sum_{\tau \subset \bar{\Omega}_i} \frac{|\tau|}{\alpha^\tau}\right)^{-1}. \tag{6.20}$$

In the case of constant α in Ω_i the (local) basis functions simplify to be given at $x \in \Omega_i$ by

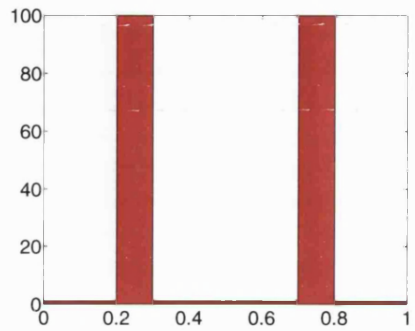
$$\vartheta^{Ms}_{i-1,\Omega_i}(x) = 1 - (x - (i-1)H)/H$$

and by

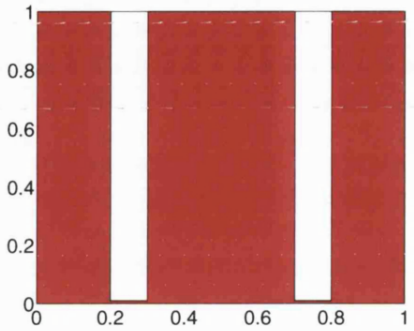
$$\vartheta^{Ms}_{i,\Omega_i}(x) = (x - (i-1)H)/H,$$

i.e. the linear coarse grid basis functions.

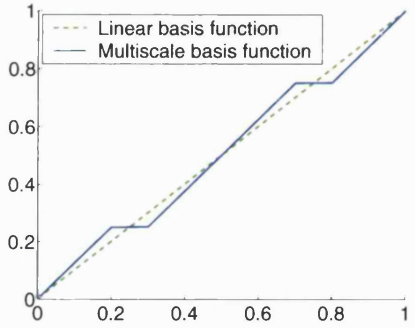
The dependence of the multiscale basis functions on α can also be seen in Figure 6-3 for a simple two media problem in which we set $\alpha = 1$ in parts of Ω and $\alpha = \hat{\alpha}$ in other parts and where $\hat{\alpha} \to \infty$, resp. $\hat{\alpha} \to 0$.



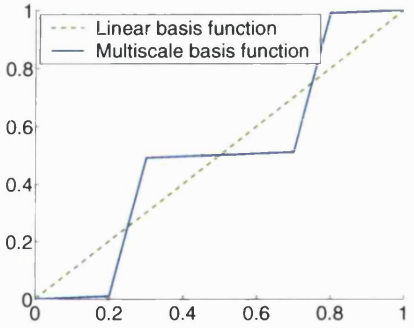
(a) Two large coefficients



(b) Two small coefficients



(c) Basis functions for two large coefficients



(d) Basis functions for two small coefficients

Figure 6-3: 1D Linear and multiscale basis functions on $\Omega_i = [(i-1)H, iH]$, where on the horizontal axis t is plotted, such that $x = x_{i-1} + tH$ and on the vertical axis $\psi_i^{Lin}(x)$, resp. $\psi_i^{Ms}(x)$ is plotted

6.1.2.2 Higher dimensional case

Now consider a domain $\Omega \subset \mathbb{R}^d$, $d \geq 2$, and a triangulation (i.e. triangular elements in 2D, tetrahedral elements in 3D) of the domain. A local multiscale basis function $\vartheta_{i,\bar{\Omega}_j}^{Ms}$ for every coarse grid element $\bar{\Omega}_j$ and every coarse grid node $\mathbf{x}_i^H \in \bar{\Omega}_j$ and from these a global multiscale basis functions ψ_i^{Ms} for every coarse grid node $\mathbf{x}_i^H \in \mathcal{N}^H(\Omega)$ are computed as in (6.3) and (6.4).

However the choice of boundary conditions that we impose on $\vartheta^{Ms}_{i,\Omega_j}$ for $j = 1, ..., p$ and $\mathbf{x}^H_i \in \bar{\Omega}_j$, when solving (6.3) is more complicated now. We therefore next want to specify these boundary conditions. As before, we would like the local basis functions $\vartheta^{Ms}_{i,\Omega_j}$ to satisfy $\vartheta^{Ms}_{i,\Omega_j}(\mathbf{x}^H_k) = \delta_{i,k}$ for any coarse grid nodes $\mathbf{x}^H_i$ and $\mathbf{x}^H_k$ in $\bar{\Omega}_j$. However, unlike in the one-dimensional case, this does not define the boundary conditions completely as there are fine grid nodes on the coarse grid edges (and faces for $d = 3$) on which we can choose the boundary conditions in many different ways.
But what is the ideal choice of boundary conditions, if we use the multiscale basis functions for the coarsening in our domain decomposition preconditioners? And in what sense can we call boundary conditions ideal then?

6.1.3 Boundary conditions

The simplest choice of boundary conditions for the construction of the multiscale basis functions are the obvious *linear boundary conditions*, which correspond exactly to the linear basis functions on the subdomain boundaries. (Using these boundary conditions, (6.3) and (6.4) the multiscale basis functions are in fact identical to the linear coarse grid basis functions inside a subdomain Ω_j if α is constant on Ω_j). These boundary conditions are easy to construct, but do not always perform very well, as the numerical tests at the end of this chapter will show.

With a little bit more computational effort, we can also construct the so called *oscillatory boundary conditions*. Let us first of all consider the 2D case and assume we want to construct the local basis function $\vartheta^{Ms}_{i,\Omega_j}$ for $\mathbf{x}^H_i \in \bar{\Omega}_j$, on subdomain Ω_j (with boundary Γ_j and with corner nodes $\mathbf{x}^H_i$, $\mathbf{x}^H_k$ and $\mathbf{x}^H_l$). Furthermore let $\mu_{i,j} := \vartheta^{Ms}_{i,\Omega_j}|_{\Gamma_j}$. We can then construct μ_{ij} by solving one dimensional problems on the edges of Ω_j. For example, let $E_{i,k}$ be the (open) edge with end nodes $\mathbf{x}^H_i, \mathbf{x}^H_k \in \Gamma_j$ and let $E_{i,\mathbf{x}}$ be the (open) edge with end nodes $\mathbf{x}^H_i, \mathbf{x} \in \Gamma_j$ (see Figure 6-1). Furthermore define for each $\mathbf{x} \in \Gamma_j$

$$\widetilde{\alpha}(\mathbf{x}) := \max_{\tau:\mathbf{x}\in\tau} (\alpha^\tau). \tag{6.21}$$

The maximum of the permeability α in the neighbouring elements is chosen in (6.21) since the maximum flow is through areas with large permeability. Experiments for two-media problems and for random permeability fields show furthermore, that this choice of $\widetilde{\alpha}$ is in fact favourable. Let $\int_{E_{k,l}}$ denote a (line) integral over edge $E_{k,l}$. Then on Γ_j the oscillatory boundary condition $\mu_{i,j}$ for the local multiscale basis function $\vartheta^{Ms}_{i,\Omega_j}$ on

the edges $E_{i,k}$, $E_{i,l}$ and $E_{k,l}$ is given by

$$\mu_{i,j}(\mathbf{x}) := 1 - \left(\int_{E_{i,\mathbf{x}}} \frac{1}{\widetilde{\alpha}(\mathbf{t})} \mathrm{d}\mathbf{t} \right) \left(\int_{E_{i,k}} \frac{1}{\widetilde{\alpha}(\mathbf{t})} \mathrm{d}\mathbf{t} \right)^{-1} \quad \text{for } \mathbf{x} \in E_{i,k}, \tag{6.22}$$

$$\mu_{i,j}(\mathbf{x}) := 1 - \left(\int_{E_{i,\mathbf{x}}} \frac{1}{\widetilde{\alpha}(\mathbf{t})} \mathrm{d}\mathbf{t} \right) \left(\int_{E_{i,l}} \frac{1}{\widetilde{\alpha}(\mathbf{t})} \mathrm{d}\mathbf{t} \right)^{-1} \quad \text{for } \mathbf{x} \in E_{i,l}, \tag{6.23}$$

$$\mu_{i,j}(\mathbf{x}) := 0 \text{ for } \mathbf{x} \in E_{k,l} \text{ and} \tag{6.24}$$

$$\mu_{i,j}(\mathbf{x}_i^H) := 1. \tag{6.25}$$

(These boundary conditions are solutions of one-dimensional problems similar to system (6.9) with α replaced by $\widetilde{\alpha}$ and the domain $(0,1)$ replaced by the edge on which we describe the boundary condition. If α is chosen constant on each fine grid element, then these functions are piecewise linear on the fine grid edges.)

In practice these functions only have to be sampled at the nodes of the fine mesh on each edge of Γ_j.
Now consider two neighbouring subdomains Ω_j and Ω_m and a coarse grid node $\mathbf{x}_i^H$ on the common boundary between these subdomains. Then we know by the definition (6.21) of $\widetilde{\alpha}$ that for all points $\mathbf{x}$ on this common boundary

$$\mu_{i,j}(\mathbf{x}) = \mu_{i,m}(\mathbf{x}). \tag{6.26}$$

Setting up the global multiscale basis functions by $\psi_i^{Ms}|_{\Omega_j} := \vartheta_{i,\Omega_j}^{Ms}$, we can therefore see that these global basis functions are continuous on the whole of Ω, and therefore H^1-conforming.

While linear boundary conditions do not take the underlying field into consideration on Γ_j, the oscillatory boundary conditions are based on the permeability field α. As we have a direct formulation for them, their setup time does almost not differ from that of the linear boundary conditions. However, as numerical results in this and in the next chapter will show, their performance is sometimes clearly superior to that of the linear boundary conditions.

In the 3D case we can construct oscillatory boundary conditions $\mu_{i,j}$ for the local coarse grid multiscale basis function $\vartheta_{i,\Omega_j}^{Ms}$ on a tetrahedral subdomain Ω_j in a three

step process:

(i) Fix the values of $\mu_{i,j}$ on the coarse grid nodes such that $\mu_{i,j}\left(\mathbf{x}_i^H\right) = 1$ and $\mu_{i,j}\left(\mathbf{x}_k^H\right) = 0$ for all other coarse grid nodes $\mathbf{x}_k^H \in \mathcal{N}^H(\Omega)$.

(ii) Define for all $\mathbf{x} \in \Gamma_j$, $\widetilde{\alpha}(\mathbf{x}) := \max_{\tau:\mathbf{x}\in\tau}(\alpha^\tau)$ (remembering that the elements $\tau \in \mathcal{T}^h(\bar{\Omega})$ are closed). With this solve 1D problems on the edges of $\bar{\Omega}_j$ exactly as for the construction of the boundary conditions in the 2D case with the boundary conditions given in (i), resp. use similar formulae as given in the 2D case.

(iii) Using the values from the 1D problems at the coarse element edges, the values from (i) at the coarse grid nodes and the coefficients $\widetilde{\alpha}(\mathbf{x}) := \max_{\tau:\mathbf{x}\in\tau}(\alpha^\tau)$ for all $\mathbf{x}$ on a given coarse grid face, solve 2D problems on the faces of the subdomain. (We therefore can use the grid of triangles given on the face by the original tetrahedral fine grid elements and introduce a new coordinate system in the plane of the face).

6.1.4 Implementation aspects of the preconditioner

Having defined the global multiscale basis functions ψ_i^{Ms} for $\mathbf{x}_i^H \in \mathcal{N}^H(\Omega)$, the $n_h \times n_H$ *multiscale extension matrix*

$$R_{0,Ms}^T := [\psi_1^{Ms}, ..., \psi_{n_H}^{Ms}] \tag{6.27}$$

(where $\boldsymbol{\psi}_i^{Ms}$, $i = 1, ..., p$, are the vectors of values of ψ_i^{Ms} at the fine grid nodes in Ω), the coarse grid stiffness matrix

$$A_{0,Ms} := R_{0,Ms} A R_{0,Ms}^T, \tag{6.28}$$

and the *two-level additive Schwarz preconditioner* with multiscale coarsening

$$M_{2,Ms}^{-1} := \sum_{j=1}^{p} R_j^T A_j^{-1} R_j + R_{0,Ms}^T A_{0,Ms}^{-1} R_{0,Ms} = M_1^{-1} + R_{0,Ms}^T A_{0,Ms}^{-1} R_{0,Ms}, \tag{6.29}$$

we can now take a first look at how to implement this new preconditioner in an effective way:

- The contributions of the multiscale basis functions ψ_i^{Ms} can be computed completely independently for the different subdomains. Therefore this computation can be almost perfectly parallelised.

- Since $A_j := R_j A R_j^T$, $j = 1, ..., p$, and $A_{0,Ms} := R_{0,Ms} A R_{0,Ms}^T$ and since for our preconditioners we only solve problems with the stiffness matrices A_j, $j = 1, ..., p$, we never need the matrices R_j, $j = 1, ..., p$, and $R_{0,Ms}$ on their own and do not set them up explicitly for our computations. Since A, as well as R_j, $j = 1, ..., p$, and $R_{0,Ms}$ are sparse so will be the matrices A_j, $j = 1, ..., p$, and $A_{0,Ms}$ (with similar sparsity patterns as A). Considering these two aspects, we can reduce the storage of the preconditioner significantly.

- Due to the nature of the decomposition either A_j or $A_{0,Ms}$ will be of size $\geq \sqrt{n}$ where n is the dimension of the global stiffness matrix A. It is therefore for large n sometimes expensive to factorise either A_j or $A_{0,Ms}$ or both of them exactly (even when using direct methods, that make use of the sparsity patterns of the matrices).

6.2 Convergence theory

We now would like to find out, if and how MsFE-coarsening can improve the convergence of the preconditioned cg-method using additive Schwarz preconditioners. Let us therefore first of all consider the following special case for which we can give sharp theoretical bounds and then give numerical results for the general case in the next subsection.

6.2.1 Special case

Assumption A: *$\alpha(\mathbf{x})$ is constant and equal to 1 in the overlaps of any subdomains (see Figure 6-4) and $\alpha(\mathbf{x}) > 1$ only for $\mathbf{x} \in \Omega$ outside the overlaps. In the rest of this section when discussing multiscale basis functions, we assume that linear boundary conditions on each Ω_j are used (see Subsection 6.1.3).*
Additionally to this assume that $\alpha(\mathbf{x}) \geq 1$, $\mathbf{x} \in \Omega$, which (when writing the norms in elementwise form) directly leads to

$$\|v\|_{L_2(\Omega)} \leq \|v\|_{L_2(\Omega),\alpha}, \; v \in L_2(\Omega), \; \text{and} \; \|v\|_{H^1(\Omega)} \leq \|v\|_{H^1(\Omega),\alpha}, \; v \in H^1(\Omega). \quad (6.30)$$

For $\mathbf{x}_i^H \in \mathcal{N}^H(\Omega)$ let ψ_i^0 denote the coarse grid basis functions of $\mathcal{V}_0$ as defined in (5.17) (where $\psi_i^0 := \psi_i^{Lin}$ in the case of linear and $\psi_i^0 := \psi_i^{Ms}$ in the case of MsFE-interpolation). In each case the ψ_i^0 will be piecewise linear with respect to the fine

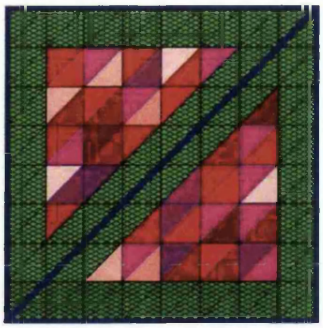


Figure 6-4: Example for subdomains considered in the convergence analysis (green=overlap; in the overlap: $\alpha^\tau = 1$; in areas with shades of red: $\alpha^\tau \geq 1$)

mesh, but will have freedoms only at the coarse grid nodes $\mathbf{x}_i^H \in \mathcal{N}^H(\Omega)$. Furthermore let

$$\omega_{\Omega_j} := \text{interior}\left(\bigcup_{k:\bar{\Omega}_k \in \mathcal{T}^H(\bar{\Omega}), \bar{\Omega}_k \cap \bar{\Omega}_j \neq \emptyset} \bar{\Omega}_k\right) \tag{6.31}$$

(as originally defined in (5.16)).
Let

$$\widetilde{\omega_{\Omega_j}} := \begin{cases} \omega_{\Omega_j}, & \text{if } \omega_{\Omega_j} \text{ does not touch } \Gamma = \partial\Omega \text{ or touches } \Gamma \text{ in at least one edge (for d=2) or at least one face (for d=3).} \\ \text{interior}\left(\bar{\omega}_{\Omega_j} \cup \bar{\Omega}_k\right), & \text{otherwise, where } \Omega_k \text{ is a coarse grid element extending } \omega_{\Omega_j}, \text{ such that } \Omega_k \text{ touches } \Gamma \text{ in (at least) an edge (for d=2), resp. (at least) a face (for d=3).} \end{cases} \tag{6.32}$$

Finally let again for $\mathbf{x}_i^H \in \mathcal{N}^H(\Omega)$

$$\omega_i := \text{interior}\left(\bigcup_{j:\mathbf{x}_i^H \in \bar{\Omega}_j} \bar{\Omega}_j\right) \tag{6.33}$$

(as originally defined in (5.15)) and let $\bar{\omega}_i$ be its closure.

As in Chapter 5 define on Ω

$$\mathcal{V}^h := \mathcal{S}_0^h(\Omega),$$

where $\mathcal{S}_0^h(\Omega) := \mathcal{S}^h(\Omega) \cap H_0^1(\Omega)$ and where $\mathcal{S}^h(\Omega)$ denotes the space of continuous piecewise linear functions with respect to $\mathcal{T}^h(\bar{\Omega})$ on Ω.

Let $\mathcal{V}_0$ in this section denote the span of the coarse grid basis functions (i.e. for the examples considered in this thesis either the space of the piecewise linear function on the coarse grid or the span of the multiscale basis functions on Ω). We can then define the following indicator.

DEFINITION 6.1
Denote $\mathcal{H}_i := diam(\omega_i)$. Then we call

$$\gamma(\mathcal{V}_0, \alpha) := \max_{\mathbf{x}_i^H \in \mathcal{N}^H(\Omega)} \left\{ \mathcal{H}_i^{2-d} \left| \psi_i^0 \right|^2_{H^1(\Omega),\alpha} \right\}$$

*the **coarse space robustness indicator**.*

Remark: By the definition of the *coarse space robustness indicator*, all bounds that will be derived in this section and that depend on this indicator can be improved by minimising the energy norm $\left|\psi_i^0\right|^2_{H^1(\Omega),\alpha}$. We will furthermore see how multiscale coarsening can in fact yield a better behaved $\gamma(\mathcal{V}_0, \alpha)$ compared to linear coarsening.

EXAMPLE 6.2
Let us examine how the coarse space robustness indicator $\gamma(\mathcal{V}_0, \alpha)$ depends on α in the case of linear coarse grid basis functions $\psi_i^0 := \psi_i^{Lin}$ and in the case of MsFE coarse grid basis functions $\psi_i^0 := \psi_i^{Ms}$ for a uniform fine grid with mesh size h and uniform coarse grid with mesh size H. Here $\mathcal{H}_i = 2H$ for all i such that $\mathbf{x}_i^H \in \mathcal{N}^H(\Omega)$.

- ***Linear interpolation:***

$$\begin{aligned}
\gamma(\mathcal{V}_0, \alpha) &= \max_{\mathbf{x}_i^H \in \mathcal{N}^H(\Omega)} \left((2H)^{2-d} \left| \psi_i^{Lin} \right|^2_{H^1(\Omega),\alpha} \right) \\
&= \max_{\mathbf{x}_i^H \in \mathcal{N}^H(\Omega)} \left((2H)^{2-d} \int_\Omega \alpha \left| \nabla \psi_i^{Lin} \right|^2 \right) \\
&= 2^{3-d} \max_{\mathbf{x}_i^H \in \mathcal{N}^H(\Omega)} \left(H^{-d} \int_{\omega_i} \alpha \right)
\end{aligned}$$

We therefore obtain linear growth of the coarse space robustness indicator with $\max_{\mathbf{x}_i^H \in \mathcal{N}^H(\Omega)} \left(H^{-d} \int_{\omega_i} \alpha \right)$ in the case of linear interpolation.

- ***MsFE-interpolation:***
  Let $\Gamma_{\beta,j}$, $j \in \{1,...,p\}$ be defined as in Definition 5.12. Recall that by Assumption A, $\alpha(\mathbf{x}) = 1$, for all points $\mathbf{x} \in \Gamma_{\beta,j}$. In this case the oscillatory boundary conditions of Subsection 6.1.3 will reduce to linear boundary conditions and so from now on we can assume linear boundary conditions. Finally suppose:

  Assumption B: *For any $\mathbf{x}_i^H \in \mathcal{N}^H(\Omega)$ there exists a function v_i, which is piecewise linear on the fine mesh, which vanishes on $\Omega_j \backslash \Gamma_{\beta,j}$ for all $\Omega_j \subset \omega_i$, which is equal to ψ_i^{Lin} on Γ_j, $j = 1,...,p$, and for which there exists a constant C, such that then $|v_i|^2_{H^1(\Gamma_{\beta,j})} \leq C(\beta H)^{-2} |\Gamma_{\beta,j}|$.*
  (After the following proof we will show that such a function v_i really exists.)

  Since with

$$a(u,v) := \int_{\Omega_j} \alpha \nabla u . \nabla v, \qquad \forall u, v \in H^1(\Omega_j),$$

  and since by the definitions of ψ_i^{Ms} and v_i and by integration by parts we obtain $a(\psi_i^{Ms}, \psi_i^{Ms} - v_i) = 0$, it follows that on each subdomain Ω_j

$$\int_{\Omega_j} \alpha \left|\nabla \psi_i^{Ms}\right|^2 \leq \int_{\Omega_j} \alpha \left|\nabla v_i\right|^2 . \tag{6.34}$$

  Therefore

$$0 \leq a(v_i - \psi_i^{Ms}, v_i - \psi_i^{Ms}) = a(v_i, v_i) - a(v_i, \psi_i^{Ms}) = a(v_i, v_i) - a(\psi_i^{Ms}, \psi_i^{Ms}) \tag{6.35}$$

  and (6.34) follows.

Using this minimisation bound, we obtain

$$
\begin{aligned}
\gamma(\mathcal{V}_0,\alpha) &= \max_{\mathbf{x}_i^H\in\mathcal{N}^H(\Omega)}\left((2H)^{2-d}\left|\psi_i^{Ms}\right|^2_{H^1(\Omega),\alpha}\right)\\
&= \max_{\mathbf{x}_i^H\in\mathcal{N}^H(\Omega)}\left((2H)^{2-d}\int_\Omega \alpha\left|\nabla\psi_i^{Ms}\right|^2\right)\\
&= \max_{\mathbf{x}_i^H\in\mathcal{N}^H(\Omega)}\left((2H)^{2-d}\int_{\omega_i} \alpha\left|\nabla\psi_i^{Ms}\right|^2\right)\\
&= \max_{\mathbf{x}_i^H\in\mathcal{N}^H(\Omega)}\left((2H)^{2-d}\sum_{j:\Omega_j\subset\omega_i}\left(\int_{\Omega_j} \alpha\left|\nabla\psi_i^{Ms}\right|^2\right)\right)\\
&\leq \max_{\mathbf{x}_i^H\in\mathcal{N}^H(\Omega)}\left((2H)^{2-d}\sum_{j:\Omega_j\subset\omega_i}\left(\int_{\Omega_j} \alpha\left|\nabla v_i\right|^2\right)\right)\\
&= \max_{\mathbf{x}_i^H\in\mathcal{N}^H(\Omega)}\left((2H)^{2-d}\sum_{j:\Omega_j\subset\omega_i}\left(\int_{\Gamma_{\beta,j}} \alpha\left|\nabla v_i\right|^2\right)\right)\\
&\quad (\text{ since } \nabla v_i(\mathbf{x}) = 0 \text{ for } \mathbf{x}\in\Omega_j\backslash\Gamma_{\beta,j}\ )\\
&= \max_{\mathbf{x}_i^H\in\mathcal{N}^H(\Omega)}\left((2H)^{2-d}\sum_{j:\Omega_j\subset\omega_i}\left(\int_{\Gamma_{\beta,j}} \left|\nabla v_i\right|^2\right)\right)\\
&\leq C\max_{\mathbf{x}_i^H\in\mathcal{N}^H(\Omega)}\left((2H)^{2-d}\sum_{j:\Omega_j\subset\omega_i}(\beta H)^{-2}\left|\Gamma_{\beta,j}\right|\right)\\
&\quad (\text{ by the definition of } v_i\ )\\
&\leq C\beta^{-1}\\
&\quad (\text{ since } |\Gamma_{\beta,j}|\leq C\beta H^d\ )
\end{aligned}
$$

for a generic constant C (independent of α, β, h and H). Therefore the coarse space robustness indicator is independent of α for multiscale basis functions for this model problem with $\alpha = 1$ everywhere in the overlap of any subdomains.

Construction of v_i: *We now have to show that a function v_i that satisfies Assumption B really exists. Let us first of all assume that each fine grid element $\tau\in\mathcal{T}^h(\bar{\Omega}_j)$ is either completely in the closure $\bar{\Gamma}_{\beta,j}$ of $\Gamma_{\beta,j}$ or completely in the closure of $\Omega_j\backslash\Gamma_{\beta,j}$. Divide each subdomain Ω_j into three regions:*

- ***Region** $R_{j,1}$: Given simply by $\Omega_j\backslash\Gamma_{\beta,j}$, i.e. all points that are not in the overlap,*
- ***Region** $R_{j,2}$: Given by $\left(\Omega_j\backslash\Gamma_{\beta/2,j}\right)\backslash R_{j,1}$, i.e. a ring of width $\beta H/2$ around*

region $R_{j,1}$,

- **Region** $R_{j,3}$**:** *Given by* $\Omega_j \setminus (R_{j,1} \cup R_{j,2})$, *i.e. ring of width* $\beta H/2$ *around region* $R_{j,2}$.

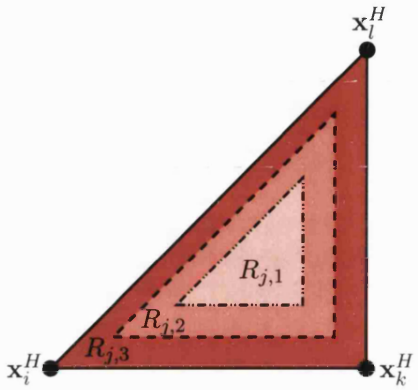


Figure 6-5: Three regions for the construction of v_i on a subdomain Ω_j

Let us now define on Ω_j *the following partition of unity*

$$\chi_B + \chi_I \equiv 1 \text{ on } \Omega_j, \tag{6.36}$$

$$\chi_I \equiv \begin{cases} 1 & \text{on } R_{j,1}, \\ 0 & \text{on } R_{j,3}, \end{cases} \tag{6.37}$$

$$\chi_B \equiv \begin{cases} 0 & \text{on } R_{j,1}, \\ 1 & \text{on } R_{j,3}. \end{cases} \tag{6.38}$$

By Lemma 5.6 we can furthermore choose χ_B *and* χ_I *in such a way, that there exists a (generic) constant* C *independent of* β *and* H *such that*

$$|\nabla \chi_B(\mathbf{x})| \leq C\beta^{-1}H^{-1} \quad \text{and} \quad |\nabla \chi_I(\mathbf{x})| \leq C\beta^{-1}H^{-1}, \ \mathbf{x} \in \Omega. \tag{6.39}$$

Now if ψ_i^{Lin} *is the linear coarse grid basis function, which is* 1 *at* $\mathbf{x}_i^H \in \mathcal{N}^H(\Omega)$ *and* 0 *at all other coarse grid nodes, we can set for* $\mathbf{x} \in \Omega_j$

$$v_i(\mathbf{x}) := I^h\left(\chi_B \psi_i^{Lin}(\mathbf{x})\right), \tag{6.40}$$

where I^h *here is the finite element interpolation onto the finite element space*

$\mathcal{V}_j := \left\{ v_h \in \mathcal{V}^h : \ supp\,(v_h) \subset \widetilde{\Omega_j} \right\}$ *(as originally defined in Section 5.1).*
By this we obtain the following properties for v_i

(i) v_i *is in the desired finite element space, i.e. piecewise linear on* $T^h(\bar{\Omega})$,

(ii) $v_i(\mathbf{x}) = 0$ *for* $\mathbf{x} \in \Omega_j \backslash \Gamma_{\beta,j}$,

(iii) $v_i|_{\Gamma_j} = \psi_i^{Lin}|_{\Gamma_j}$

(iv) $|v_i|^2_{H^1(\Gamma_{\beta,j})} \leq C(\beta H)^{-2} \left|\Gamma_{\beta,j}\right|$.

The first three required properties of v_i, *i.e. (i), (ii) and (iii), follow immediately. To prove (iv) let* $\hat{\chi}_{B,\tau}$ *be the average value of* χ_B *on element* $\tau \subset \bar{\Omega}_j$. *Then by*

$$\begin{aligned} \boldsymbol{\nabla} v_i(\mathbf{x}) &= \boldsymbol{\nabla} I^h \left(\chi_B \psi_i^{Lin}\right)(\mathbf{x}) \\ &= \boldsymbol{\nabla} I^h \left(\left(\chi_B - \hat{\chi}_{B,\tau}\right) \psi_i^{Lin}\right)(\mathbf{x}) + \boldsymbol{\nabla} \left(\hat{\chi}_{B,\tau} \psi_i^{Lin}(\mathbf{x})\right). \end{aligned} \tag{6.41}$$

Hence, we obtain

$$|v_i|^2_{H^1(\tau)} \leq 2 \left| I^h \left(\chi_B - \hat{\chi}_{B,\tau}\right) \psi_i^{Lin} \right|^2_{H^1(\tau)} + 2 \left| \hat{\chi}_{B,\tau} \psi_i^{Lin} \right|^2_{H^1(\tau)}. \tag{6.42}$$

Now using the inverse inequality $|v|_{H^1(\tau)} \leq Ch^{-1} \|v\|_{L_2(\tau)}$ *for any linear function* v *on* τ *and* $\left\|I^h(f.v)\right\|_{L_2(\tau)} \leq \|f\|_{L_\infty(\tau)} \|v\|_{L_2(\tau)}$ *for any continuous function* f *and for any linear function* v *on* τ, *we obtain*

$$\begin{aligned} \left| I^h \left(\left(\chi_B - \chi_{B,\tau}\right) \psi_i^{Lin}\right) \right|_{H^1(\tau)} &\leq Ch^{-1} \left\| I^h \left(\left(\chi_B - \hat{\chi}_{B,\tau}\right) \psi_i^{Lin}\right) \right\|_{L_2(\tau)} \\ &\leq Ch^{-1} \left\|\chi_B - \chi_{B,\tau}\right\|_{L_\infty(\tau)} \left\|\psi_i^{Lin}\right\|_{L_2(\tau)} \\ &\leq C\beta^{-1} H^{-1} \left|\tau\right|^{1/2}. \end{aligned} \tag{6.43}$$

Moreover

$$\left|\hat{\chi}_{B,\tau} \psi_i^{Lin}\right|_{H^1(\tau)} = \left|\hat{\chi}_{B,\tau}\right| \left|\psi_i^{Lin}\right|_{H^1(\tau)} \leq \left|\psi_i^{Lin}\right|_{H^1(\tau)} \leq H^{-1} \left|\tau\right|^{1/2}. \tag{6.44}$$

Combining (6.42), (6.43) and (6.44), we get

$$\begin{aligned} |v_i|^2_{H^1(\Gamma_{\beta,j})} &= \sum_{\tau \in \Gamma_{\beta,j}} |v_i|^2_{H^1(\tau)} \\ &\leq \sum_{\tau \in \Gamma_{\beta,j}} \left(2 \left| I^h \left(\chi_B - \hat{\chi}_{B,\tau}\right) \psi_i^{Lin} \right|^2_{H^1(\tau)} + 2 \left| \hat{\chi}_{B,\tau} \psi_i^{Lin} \right|^2_{H^1(\tau)} \right) \\ &\leq \sum_{\tau \in \Gamma_{\beta,j}} \left(C\beta^{-2} H^{-2} \left|\tau\right| + 2H^{-2} \left|\tau\right| \right) \\ &\leq C\beta^{-2} H^{-2} \left|\Gamma_{\beta,j}\right|, \end{aligned} \tag{6.45}$$

where C is a constant. This finishes the proof of (iv).

We can then prove the following bounds similarly to classical bounds given in Toselli and Widlund [128], Lemma 3.6.

LEMMA 6.3
If Assumption A holds, then there exists a constant C (independent of α, β, h and H) and a linear operator $\bar{I}^H : H_0^1(\Omega) \rightarrow \mathcal{V}^h$ such that for all $u \in H_0^1(\Omega)$ and for all $\bar{\Omega}_j \in \mathcal{T}^H(\bar{\Omega})$

$$\left\|u - \bar{I}^H u\right\|^2_{L_2(\Omega_j)} \leq CH^2 \left|u\right|^2_{H^1\left(\widetilde{\omega_{\Omega_j}}\right)}, \tag{6.46}$$

$$\left|\bar{I}^H u\right|^2_{H^1(\Omega_j),\alpha} \leq C\gamma\left(\mathcal{V}_0, \alpha\right) \left|u\right|^2_{H^1\left(\widetilde{\omega_{\Omega_j}}\right),\alpha}. \tag{6.47}$$

Proof:
Consider a standard quasi-interpolant $\bar{I}^H$ such that

$$\bar{I}^H u := \sum_{\mathbf{x}_i^H \in \mathcal{N}^H(\Omega)} \bar{u}_i \psi_i^0, \tag{6.48}$$

where

$$\bar{u}_i := \left|\omega_i\right|^{-1} \int_{\omega_i} u \tag{6.49}$$

and where ψ_i^0, $\mathbf{x}_i^H \in \mathcal{N}^H(\Omega)$, are either the multiscale basis functions ψ_i^{Ms} or the linear basis functions ψ_i^{Lin} on Ω.
For the multiscale basis functions we now would like to find an upper bound for $\left\|\psi_i^{Ms}\right\|_{L_\infty(\Omega)}$. It follows by the discrete maximum principle (see e.g. inequality (5) of Jüngel and Unterreiter [80]), that for all i with $\mathbf{x}_i^H \in \mathcal{N}^H(\Omega)$

$$\inf_{\mathbf{x}\in\Omega} \psi_i^{Ms}(\mathbf{x}) \geq \inf_j \left(\inf_{\mathbf{x}\in\bar{\Omega}_j : \bar{\Omega}_j \in \mathcal{T}^H(\bar{\Omega})} \psi_i^{Ms}|_{\bar{\Omega}_j} \right) \geq \inf_j \left(\inf_{\mathbf{x}\in\Gamma_j : \bar{\Omega}_j \in \mathcal{T}^H(\bar{\Omega})} \psi_i^{Ms}|_{\Gamma_j} \right) \tag{6.50}$$

and e.g. by Proposition 11 of Jüngel and Unterreiter [80], that

$$\sup_{\mathbf{x}\in\Omega} \psi_i^{Ms}(\mathbf{x}) \leq \sup_j \left(\sup_{\mathbf{x}\in\bar{\Omega}_j : \Omega_j \in \mathcal{T}^H(\bar{\Omega})} \psi_i^{Ms}|_{\bar{\Omega}_j} \right) \leq \sup_j \left(\sup_{\mathbf{x}\in\Gamma_j : \bar{\Omega}_j \in \mathcal{T}^H(\bar{\Omega})} \psi_i^{Ms}|_{\Gamma_j} \right). \tag{6.51}$$

In the example considered (with $\alpha = 1$ in every overlap) the oscillatory boundary conditions are identical to the linear boundary conditions for the construction of the

multiscale basis functions. For this example (and by equations (6.22)-(6.24) if $\alpha(\mathbf{x}) \geq 0$, $\mathbf{x} \in \Omega$, in fact also for oscillatory boundary conditions for general fields) it follows that for all $j \in \{1, ..., p\}$, and all i with $\mathbf{x}_i^H \in \mathcal{N}^H(\Omega)$,

$$0 \leq \psi_i^{Ms}|_{\Gamma_j} \leq 1. \tag{6.52}$$

Combining (6.52) with (6.50) and (6.51) it follows that

$$0 \leq \inf_{\Omega} \psi_i^{Ms} \leq \sup_{\Omega} \psi_i^{Ms} \leq 1 \tag{6.53}$$

and therefore that

$$\left\|\psi_i^{Ms}\right\|_{L_\infty(\Omega)} \leq 1. \tag{6.54}$$

For linear coarse grid functions, we also know $\left\|\psi_i^{Lin}\right\|_{L_\infty(\Omega)} \leq 1$, and combining both cases

$$\left\|\psi_i^{0}\right\|_{L_\infty(\Omega)} \leq 1. \tag{6.55}$$

Furthermore due to the quasi-uniformity of the coarse mesh, this implies (see Lemma B.5 in Toselli and Widlund [128]) that there exists a constant C (independent of α, β, h and H) such that

$$\left\|\psi_i^{0}\right\|_{L_2(\Omega_j)} \leq CH^{d/2} \tag{6.56}$$

and

$$\begin{aligned} |\bar{u}_i| &= \left| |\omega_i|^{-1} \int_{\omega_i} u\left(\mathbf{x}\right) d\mathbf{x} \right| \\ &\leq |\omega_i|^{-1} \left| \int_{\omega_i} u\left(\mathbf{x}\right) d\mathbf{x} \right| \\ &\leq |\omega_i|^{-1} \left\{ \int_{\omega_i} 1^2 \right\}^{1/2} \|u\|_{L_2(\omega_i)} \\ &= |\omega_i|^{-1/2} \|u\|_{L_2(\omega_i)} \\ &\leq CH^{-d/2} \|u\|_{L_2(\omega_i)}. \end{aligned} \tag{6.57}$$

We therefore obtain

$$\left\|\bar{I}^H u\right\|_{L_2(\Omega_j)} \leq \sum_{i:\mathbf{x}_i^H \in \Omega_j} \left(CH^{-d/2} \|u\|_{L^2\left(\omega_{\Omega_j}\right)} \right) \left\|\psi_i^0\right\|_{L_2\left(\omega_{\Omega_j}\right)} \leq C \|u\|_{L_2\left(\omega_{\Omega_j}\right)}. \tag{6.58}$$

For the triangulation $\mathcal{T}^H(\bar{\Omega})$ of the domain Ω into subdomains Ω_j we now have to distinguish three different cases:

- **Case 1:** ω_{Ω_j} does not touch Γ.
- **Case 2:** ω_{Ω_j} touches Γ in at least an edge (for $d = 2$), resp. a face (for $d = 3$).
- **Case 3:** ω_{Ω_j} touches Γ only in nodes (for $d = 2$) or only in nodes and/or edges (for $d = 3$).

We get for

- **Case 1:**
  Let now
  $$\hat{u} := u - \left|\omega_{\Omega_j}\right|^{-1} \int_{\omega_{\Omega_j}} u. \tag{6.59}$$
  We know by Poincaré's inequality (see Toselli and Widlund [128], Corollary A.15) that for any subdomain Ω_j (with diameter H), there exists a constant C, that depends only on the shape of ω_{Ω_j} but not on its size, such that
  $$\|\hat{u}\|^2_{L_2\left(\omega_{\Omega_j}\right)} \leq CH^2 \left|\hat{u}\right|^2_{H^1\left(\omega_{\Omega_j}\right)}. \tag{6.60}$$
  Using this inequality and since, if $\mathcal{N}^H\left(\bar{\Omega}_j\right)$ is the set of coarse grid nodes of subdomain Ω_j (including the nodes on Γ_j),
  $$\sum_{\mathbf{x}_i^H \in \mathcal{N}^H\left(\bar{\Omega}_j\right)} \psi_i^0\left(\mathbf{x}\right) = 1, \tag{6.61}$$
  and therefore $\bar{I}^H$ reproduces constant functions on Ω_j, it follows using (6.58) and (6.60) that
  $$\left\|u - \bar{I}^H u\right\|_{L_2(\Omega_j)} = \left\|\hat{u} - \bar{I}^H \hat{u}\right\|_{L_2(\Omega_j)} \leq C \left\|\hat{u}\right\|_{L_2\left(\omega_{\Omega_j}\right)} \leq CH^2 \left|u\right|_{H^1\left(\omega_{\Omega_j}\right)},$$
  since ω_{Ω_j} has diameter of order H and since $|\hat{u}|_{H^1(\omega_{\Omega_j})} = |u|_{H^1(\omega_{\Omega_j})}$. Hence we have proved (6.46) for this case.
- **Case 2:**
  Now consider the case in which ω_{Ω_j} touches Γ in at least an edge (for $d = 2$), resp. a face (for $d = 3$). We then get using Friedrich's inequality (see Lemma 5.11), since $u \in H_0^1(\Omega)$,
  $$\|u\|^2_{L_2(\omega_{\Omega_j})} \leq CH^2 \left|u\right|^2_{H^1(\omega_{\Omega_j})}. \tag{6.62}$$

Combining this inequality with (6.58), we then obtain

$$\left\|\bar{I}^H u\right\|_{L_2(\Omega_j)} \leq C \left\|u\right\|_{L_2\left(\omega_{\Omega_j}\right)} \leq CH^2 \left|u\right|^2_{H^1(\omega_{\Omega_j})}, \tag{6.63}$$

which again finishes the proof of (6.46).

- **Case 3:**
  Finally consider the case in which ω_{Ω_j} touches Γ only in nodes (for $d = 2$), resp. only in edges and/or nodes (for $d = 3$). Then we add one coarse grid element $\bar{\Omega}_k \in \mathcal{T}^H(\bar{\Omega})$, such that $\widetilde{\omega_{\Omega_j}} :=$ interior $\left(\bar{\omega}_{\Omega_j} \cup \bar{\Omega}_k\right)$ (as defined in (6.32)) has got (at least) one common edge (for $d = 2$), resp. (at least) one common face with Γ (for $d = 3$).
  We can then reduce this case to Case 2 if we replace ω_{Ω_j} by $\widetilde{\omega_{\Omega_j}}$.

To prove inequality (6.47) note again that for any node $\mathbf{x}_i^H \in \mathcal{N}^H(\Omega)$ inequality (6.57) holds.

We then obtain for the same three cases that we considered above:

- **Case 1:**
  Let $\hat{u} := u - \left|\omega_{\Omega_j}\right|^{-1} \int_{\omega_{\Omega_j}} u$ and let again $\mathbf{x}_i^H$ be a coarse grid node on Ω_j and so $\omega_i \subset \omega_{\Omega_j}$. Then using (6.57) we obtain

$$\begin{aligned}\left|\bar{I}^H u\right|^2_{H^1(\Omega_j),\alpha} = \left|\bar{I}^H \hat{u}\right|^2_{H^1(\Omega_j),\alpha} &\leq C \max_{\mathbf{x}_i^H \in \mathcal{N}^H(\bar{\Omega}_j)} \left(\left|\omega_i\right|^{-1} \left\|\hat{u}\right\|^2_{L_2(\omega_i)}\right) \left|\psi_i^0\right|^2_{H^1(\Omega_j),\alpha} \\ &\leq C \left|\omega_{\Omega_j}\right|^{-1} \left\|\hat{u}\right\|^2_{L_2\left(\omega_{\Omega_j}\right)} \max_{\mathbf{x}_i^H \in \mathcal{N}^H(\bar{\Omega}_j)} \left|\psi_i^0\right|^2_{H^1(\Omega_j),\alpha}.\end{aligned} \tag{6.64}$$

  Thus, since $\hat{u}$ has zero mean on ω_{Ω_j}, Poincaré's inequality (analogous to (6.60)) implies

$$\begin{aligned}\left|\bar{I}^H u\right|^2_{H^1(\Omega_j),\alpha} &\leq C \left|\omega_{\Omega_j}\right|^{-1} H^2 \left|\hat{u}\right|^2_{H^1\left(\omega_{\Omega_j}\right)} \max_{\mathbf{x}_i^H \in \mathcal{N}^H(\Omega_j)} \left|\psi_i^0\right|^2_{H^1(\Omega_j)} \\ &\leq C\gamma\left(\mathcal{V}_0, \alpha\right) \left|\hat{u}\right|^2_{H^1\left(\omega_{\Omega_j}\right)} \\ &\leq C\gamma\left(\mathcal{V}_0, \alpha\right) \left|u\right|^2_{H^1\left(\omega_{\Omega_j}\right),\alpha}.\end{aligned} \tag{6.65}$$

  For the last step we also used (6.30).

- **Case 2:**
  In this case we use Friedrich's inequality (see Lemma 5.11) again, since $u \in H_0^1(\Omega)$,

$$\|u\|^2_{L_2(\omega_{\Omega_j})} \leq CH^2 \left|u\right|^2_{H^1(\omega_{\Omega_j})}. \tag{6.66}$$

  Using this inequality together with (6.57) and inequality (6.30), we then obtain

$$\begin{aligned} \left|\bar{I}^H u\right|^2_{H^1(\Omega_j),\alpha} &\leq C \left|\omega_{\Omega_j}\right|^{-1} \|u\|^2_{L_2\left(\omega_{\Omega_j}\right)} \max_{x_i^H \in \mathcal{N}^H(\Omega)} \left|\psi_i^0\right|^2_{H^1(\Omega_j),\alpha} \\ &\leq C \left|\omega_{\Omega_j}\right|^{-1} H^2 \left|u\right|^2_{H^1\left(\omega_{\Omega_j}\right)} H^{d-2}\gamma\left(\mathcal{V}_0,\alpha\right) \\ &\leq C\gamma\left(\mathcal{V}_0,\alpha\right) \left|u\right|^2_{H^1\left(\omega_{\Omega_j}\right),\alpha}. \end{aligned} \tag{6.67}$$

- **Case 3:**
  For this final case we again add one coarse grid element $\bar{\Omega}_k \in \mathcal{T}^H(\bar{\Omega})$, such that $\widetilde{\omega_{\Omega_j}} :=$ interior $\left(\bar{\omega}_{\Omega_j} \cup \bar{\Omega}_k\right)$ (as defined in (6.32)) has got (at least) one common edge (for $d = 2$), resp. (at least) one common face with Γ (for $d = 3$).
  We can then reduce this case to Case 2 by replacing ω_{Ω_j} by $\widetilde{\omega_{\Omega_j}}$.

□

Using this lemma we can prove the following corollary for fields and overlap of the type as described above (and displayed in Figure 6-4). Let therefore the subspaces $\mathcal{V}_j \subset \mathcal{V}^h$ be defined as in Definition 5.7.

COROLLARY 6.4
If Assumption A holds, then for any $u_h \in \mathcal{V}^h$ there exist $u_i \in \mathcal{V}_i$, $i = 0, ..., p$, such that

$$u_h = \sum_{i=0}^{p} u_i \tag{6.68}$$

and

$$\sum_{i=0}^{p} |u_i|^2_{H^1(\Omega),\alpha} \leq C\left(1+\beta^{-1}\right)\gamma\left(\mathcal{V}_0,\alpha\right)|u_h|^2_{H^1(\Omega),\alpha}. \tag{6.69}$$

Proof:
Let $u_0 := \bar{I}^H u_h$. Then by Lemma 6.3 we know

$$|u_0|^2_{H^1(\Omega_j),\alpha} \leq C\cdot\gamma\left(\mathcal{V}_0,\alpha\right)|u_h|^2_{H^1\left(\widetilde{\omega_{\Omega_j}}\right),\alpha}. \tag{6.70}$$

Furthermore let $\Gamma_{\beta,j} \subset \bar{\Omega}_j$ be defined as in Definition (5.12). By the properties of the partition of unity (see Definition 5.6) and the definition of $\Gamma_{\beta,j}$ it follows for any $i = 1,...,p$, that $\chi_i(\mathbf{x}) = \delta_{ij}$ for $\mathbf{x} \in \Omega_j\backslash\Gamma_{\beta,j}$, $j = 1,...,p$. Therefore with $u_i := I^h\left(\chi_i\left(u_h - u_0\right)\right)$, where I^h is the finite element interpolation onto $\mathcal{V}^h$ (as defined in Section 5.2), we obtain

$$\begin{aligned}\sum_{i=1}^{p}|u_i|^2_{H^1(\Omega_j\backslash\Gamma_{\beta,j}),\alpha} &= |u_j|^2_{H^1(\Omega_j\backslash\Gamma_{\beta,j}),\alpha}\\ &= \left|I^h\left(\chi_j(u_h-u_0)\right)\right|^2_{H^1(\Omega_j\backslash\Gamma_{\beta,j}),\alpha}\\ &= |u_h-u_0|^2_{H^1(\Omega_j\backslash\Gamma_{\beta,j}),\alpha},\end{aligned} \tag{6.71}$$

since $\chi_j(\mathbf{x}) = 1$ for $\mathbf{x} \in \Omega_j\backslash\Gamma_{\beta,j}$ and since $u_h - u_0$ is piecewise linear on the whole of Ω with respect to the fine grid.
Now consider an element $\tau \subset \bar{\Gamma}_{\beta,j}$. With steps (5.40) to (5.46) applied exactly as in Theorem 5.7, we obtain

$$\sum_{i=1}^{p}|u_i|^2_{H^1(\tau)} \leq C\left(\beta^{-2}H^{-2}\|u_h-u_0\|^2_{L_2(\tau)} + |u_h-u_0|^2_{H^1(\tau)}\right)N_c, \tag{6.72}$$

where N_c is the number of colours as defined in Section 5.2. Since we are considering subdomains with $\alpha^\tau = 1$ for all elements τ in the overlap, this means for all $j \in \{1,...,p\}$

$$\alpha(\mathbf{x}) = 1, \text{ for all } \mathbf{x} \in \Gamma_{\beta,j} \tag{6.73}$$

and therefore for such fields and subdomains

$$\|v\|_{L_2(\Gamma_{\beta,j})} = \|v\|_{L_2(\Gamma_{\beta,j}),\alpha},\ \forall v \in L_2(\Gamma_{\beta,j}),\ \text{and}\ \|v\|_{H^1(\Gamma_{\beta,j})} = \|v\|_{H^1(\Gamma_{\beta,j}),\alpha},\ \forall v \in H^1(\Gamma_{\beta,j}). \tag{6.74}$$

Hence summing (6.72) over all $\tau \in \bar{\Gamma}_{\beta,j}$ we get

$$\begin{aligned}\sum_{i=1}^{p} |u_i|^2_{H^1(\Gamma_{\beta,j}),\alpha} &= \sum_{i=1}^{p} |u_i|^2_{H^1(\Gamma_{\beta,j})} \\ &\leq C\beta^{-2}H^{-2} \|u_h - u_0\|^2_{L_2(\Gamma_{\beta,j})} + C |u_h - u_0|^2_{H^1(\Gamma_{\beta,j})} \\ &= C\beta^{-2}H^{-2} \|u_h - u_0\|^2_{L_2(\Gamma_{\beta,j}),\alpha} + C |u_h - u_0|^2_{H^1(\Gamma_{\beta,j}),\alpha}.\end{aligned} \tag{6.75}$$

To estimate $\|u_h - u_0\|^2_{L_2(\Gamma_{\beta,j}),\alpha}$ we now use (6.74), Lemma 5.13 and inequality (6.30) to obtain

$$\begin{aligned}\|u_h - u_0\|^2_{L_2(\Gamma_{\beta,j}),\alpha} &= \|u_h - u_0\|^2_{L_2(\Gamma_{\beta,j})} \\ &\leq C\beta H^2 (1+\beta) |u_h - u_0|^2_{H^1(\Omega_j)} + C\beta \|u_h - u_0\|^2_{L_2(\Omega_j)} \\ &\leq C\beta H^2 (1+\beta) |u_h - u_0|^2_{H^1(\Omega_j),\alpha} + C\beta \|u_h - u_0\|^2_{L_2(\Omega_j)}.\end{aligned} \tag{6.76}$$

Employing (6.76) in (6.75) we get

$$\sum_{i=1}^{p} |u_i|^2_{H^1(\Gamma_{\beta,j}),\alpha} \leq C\left(1+\beta^{-1}\right) |u_h - u_0|^2_{H^1(\Omega_j),\alpha} + C\beta^{-1}H^{-2} \|u_h - u_0\|^2_{L_2(\Omega_j)}. \tag{6.77}$$

Adding (6.71) to (6.77) it then follows that

$$\sum_{i=1}^{p} |u_i|^2_{H^1(\Omega_j),\alpha} \leq C\left(1+\beta^{-1}\right) |u_h - u_0|^2_{H^1(\Omega_j),\alpha} + C\beta^{-1}H^{-2} \|u_h - u_0\|^2_{L_2(\Omega_j)}. \tag{6.78}$$

Applying the triangle inequality to the first term on the right and inequality (6.46) from Lemma 6.3 to the second term on the right we obtain

$$\sum_{i=1}^{p} |u_i|^2_{H^1(\Omega_j),\alpha} \leq C\left(1+\beta^{-1}\right) |u_h|^2_{H^1(\Omega_j),\alpha} + C\left(1+\beta^{-1}\right) |u_0|^2_{H^1(\Omega_j),\alpha} + C\beta^{-1} |u_h|^2_{H_1\left(\omega_{\Omega_j}\right)}. \tag{6.79}$$

Applying (6.70) to the second term on the right and using the fact, that since $\alpha(\mathbf{x}) \geq 1$, $\mathbf{x} \in \Omega$, we have $\|u_h\|_{H^1\left(\omega_{\Omega_j}\right)} \leq \|u_h\|_{H^1\left(\omega_{\Omega_j}\right),\alpha}$, to the third term on the right of this inequality and combining terms, we get

$$\sum_{i=1}^{p} |u_i|^2_{H^1(\Omega_j),\alpha} \leq C\left(1+\beta^{-1}\right) |u_h|^2_{H^1(\Omega_j),\alpha} + C\left(1+\beta^{-1}\right) \gamma(\mathcal{V}_0,\alpha) |u_h|^2_{H^1\left(\widetilde{\omega_{\Omega_j}}\right),\alpha}. \tag{6.80}$$

Adding (6.70) and summing over all all subdomains, it follows that

$$\sum_{i=0}^{p} |u_i|^2_{H^1(\Omega),\alpha} \leq C\left(1+\beta^{-1}\right)\gamma\left(\mathcal{V}_0,\alpha\right)|u_h|^2_{H^1(\Omega),\alpha}, \tag{6.81}$$

which finishes the proof.

□

Combining Corollary 6.4 with Theorem 5.4, we obtain

THEOREM 6.5
If Assumption A holds, then there exists a constant C (independent of α, β, h and H), such that the condition number of the preconditioned stiffness matrices using a two-level additive Schwarz preconditioner $M^{-1}_{2,AS}$ satisfies

$$\kappa\left(M^{-1}_{2,AS}A(\alpha)\right) \leq C\left(1+\beta^{-1}\right)\gamma\left(\mathcal{V}_0,\alpha\right). \tag{6.82}$$

Remarks:

- $\kappa\left(M^{-1}_{2,AS}A(\alpha)\right)$ depends on $(1+\beta^{-1})$ as predicted by the traditional theory (see Theorems 5.16 and 5.17). It also depends linearly on the coarse space robustness indicator. Improving the coarsening operator (for example by using MsFE-interpolation as shown in Example 6.2) has the potential to achieve greater robustness with respect to variations in α at least in this model problem with constant coefficients in the overlap.

- Using the bounds from Example 6.2 (i) and (ii) there exists a constant C independent of α, β, h and H for this model problem such that

$$\kappa\left(M^{-1}_{2,Lin}A(\alpha)\right) \leq C \max_{\mathbf{x}_i^H \in \mathcal{N}^H(\Omega)} \left(\int_{\omega_i} H^{-d}\alpha\right)\left(1+\beta^{-1}\right) \tag{6.83}$$

  and

$$\kappa\left(M^{-1}_{2,Ms}A(\alpha)\right) \leq C\beta^{-1}\left(1+\beta^{-1}\right). \tag{6.84}$$

Experiment A:

For a first test of the theory we choose $\alpha = \hat{\alpha}$ only in parts of the interior of the coarse grid elements (see Figure 6-6) in such a way, that $\alpha = 1$ in all overlaps of subdomains. In this case the linear and oscillatory boundary conditions are identical for the construction of the multiscale basis functions and we therefore of course expect the same iteration numbers in both cases. Now fix $h = 1/128$, $H = 1/16$, $\beta = 1/8$ and let $\alpha = \hat{\alpha}$ in two elements in the interior of the subdomains as shown in Figure 6-7.

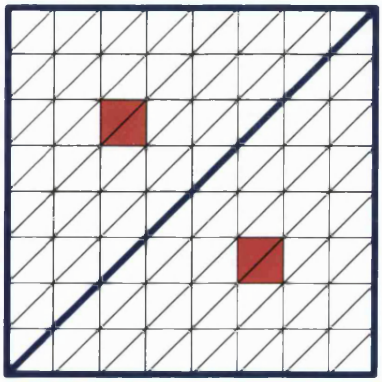


Figure 6-6: Subdomains for Experiment A (White elements: $\alpha^\tau = 1$, red elements: $\alpha^\tau = \hat{\alpha}$)

$\hat{\alpha}$	Lin. int.	Ms/Lin.b.c.	Ms/Osc.b.c.
10^{-11}	27	27	27
...	...	...	...
10^{-2}	27	27	27
10^{-1}	27	27	27
10^{0}	27	27	27
10^{1}	23	26	26
10^{2}	30	26	26
10^{3}	70	26	26
10^{5}	82	26	26
10^{7}	86	26	26
10^{9}	88	26	26

Figure 6-7: CG-iteration numbers for Experiment A for tolerance $\theta = 10^{-6}$, $h = 1/128$ and $\beta = 1/8$

As we can see in Figure 6-8, $H^{-1} \left\| \nabla \psi_i^{Ms} |_\tau \right\|_2$ is close to 0 in the two elements with $\alpha = \hat{\alpha}$, if we choose $\hat{\alpha} \gg 1$. This follows from the independence of α and the coarse space robustness indicator $\gamma(\mathcal{V}_0, \alpha)$ (see Example 6.2). The benefits of the multiscale coarsening are clearly visible in Figure 6-7. For $\hat{\alpha} \leq 1$ all interpolation types behave well. If on the other hand we choose $\hat{\alpha} \gg 1$ and use a linear coarsening operator, the number of cg-iterations grows until leveling off at about $\hat{\alpha} = 10^5$. For multiscale coarsening (with linear and with oscillatory boundary conditions) however the number of iterations stays constant for all values of $\hat{\alpha}$.

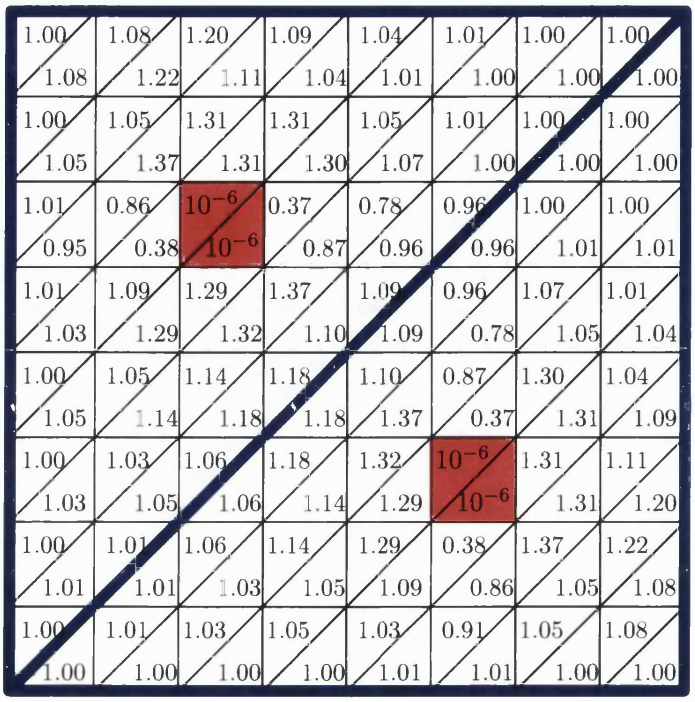


Figure 6-8: Values of $H^{-1} \cdot \left\| \nabla \psi_i^{Ms} |_\tau \right\|_2$ (multiscale interpolation) for each fine grid element τ, where ψ_i^{Ms} is the basis function which is 1 at the coarse grid node at the bottom left corner of the figure and 0 at the other coarse grid nodes, $\alpha = 10^6$ in the (red) marked elements and $\alpha = 1$ everywhere else.

6.2.2 General case

Experiment A supported the theoretical bounds developed in the previous subsection. But what happens, if we allow α also to vary within the overlaps of subdomains, i.e. the case not covered by the theory above? Will the MsFE-coarsening still work so well?

It is more complex to develop strict theoretical bounds for this more general. Details on this general theory are given in Graham et al. [60]. In this thesis however we would like to give some further numerical results, which demonstrate the power and limitations of multiscale coarsening.

Experiment B:

We will now consider a simple example for this more general case. Let us therefore move areas with permeability $\alpha = \hat{\alpha}$ to the boundary of the subdomains Ω_j. Furthermore let us specifically consider the case in which these areas consist of two elements in the middle of the coarse grid edges (as in Figure 6-9).

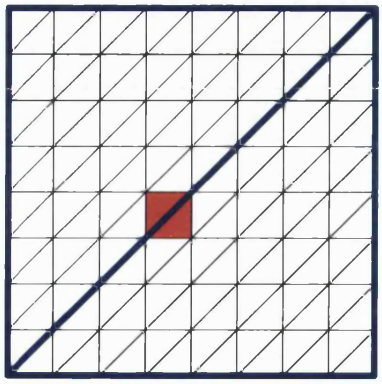


Figure 6-9: Subdomains for Experiment B (White elements: $\alpha^\tau = 1$, red elements: $\alpha^\tau = \hat{\alpha}$)

$\hat{\alpha}$	Lin. int.	Ms/Lin.b.c.	Ms/Osc.b.c.
10^{-11}	26	26	33
...	...	...	...
10^{-2}	26	26	32
10^{-1}	26	26	26
10^{0}	26	26	26
10^{1}	32	32	35
10^{2}	66	69	79
10^{3}	162	192	196
10^{5}	342	1,267	475
10^{7}	521	$> 2,000$	604
10^{9}	682	$> 2,000$	803

Figure 6-10: CG-iteration numbers for Experiment B for tolerance $\theta = 10^{-6}$, $h = 1/128$ and $\beta = 1/8$

$\hat{\alpha}$	Lin. int.	Ms/Lin.b.c.	Ms/Osc.b.c.
10^{-11}	22	22	31
...	...	...	...
10^{-2}	22	22	30
10^{-1}	22	22	23
10^{0}	22	22	22
10^{1}	23	23	23
10^{2}	25	25	23
10^{3}	51	47	23
10^{5}	61	58	23
10^{7}	67	67	23
10^{9}	67	67	23

Figure 6-11: CG-iteration numbers for Experiment B for tolerance $\theta = 10^{-6}$, $h = 1/128$ and $\beta = 1/4$

We see that for overlap $\beta H = h$, i.e. one row of elements, all preconditioners perform poorly if $\hat{\alpha} \to \infty$ (see Figure 6-10). In this case multiscale coarsening brings no advantage. However if we increase the overlap so that the whole region with $\alpha = \hat{\alpha}$ is

inside the overlap, i.e. $\beta H > h$, we obtain better results for multiscale coarsening with oscillatory boundary conditions (see Figure 6-11). In this case we can see that again for $\hat{\alpha} \leq 1$ all three types of coarsening (linear interpolation, MsFE-interpolation with linear boundary conditions and MsFE-interpolation with oscillatory boundary conditions) behave well. However the convergence of additive Schwarz preconditioners with linear interpolation and with MsFE-interpolation with linear boundary conditions for the construction of the coarse grid basis functions decreases for $\hat{\alpha} \to \infty$ up to a certain stage and then levels off. MsFE-interpolation with oscillatory boundary conditions however leads to constant iteration numbers, i.e. independence of the coefficients.
These cg-iteration numbers show that the theoretical framework developed in the beginning of this chapter is not sufficient to handle this more general case in which we allow α to be non-constant all over Ω. Unfortunately coarse space robustness is not the only factor affecting the performance of additive Schwarz methods. We also have to consider the relation between the overlap of the subdomains $\widetilde{\Omega_j}$ and the permeability field α. For this example this means that if we choose the overlap larger than the region with $\alpha = \hat{\alpha}$, the multiscale coarsening with oscillatory boundary conditions achieves a significant reduction in cg-iteration numbers compared to linear coarsening and multiscale coarsening with linear boundary conditions, while this effect cannot be observed for overlap of the size or smaller than these regions.

Experiment C:

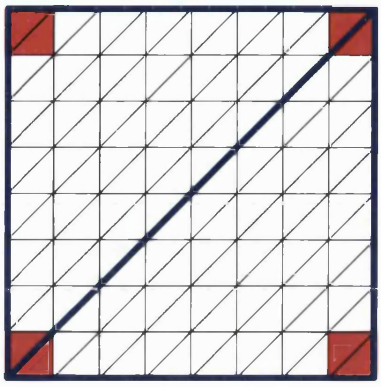


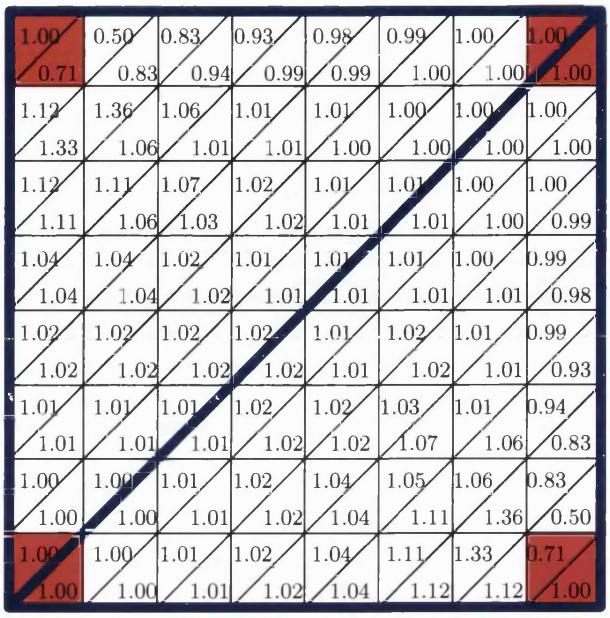
Figure 6-12: Subdomains for Experiment C (White elements: $\alpha^\tau = 1$, red elements: $\alpha^\tau = \hat{\alpha}$)

$\hat{\alpha}$	Lin. int.	Ms/Lin.b.c.	Ms/Osc.b.c.
10^{-11}	22	21	22
...	...	...	...
10^{-2}	22	21	22
10^{-1}	22	21	22
10^{0}	22	22	22
10^{1}	26	22	22
10^{2}	60	39	23
10^{3}	90	79	23
10^{5}	98	97	26
10^{7}	110	108	30
10^{9}	116	115	33

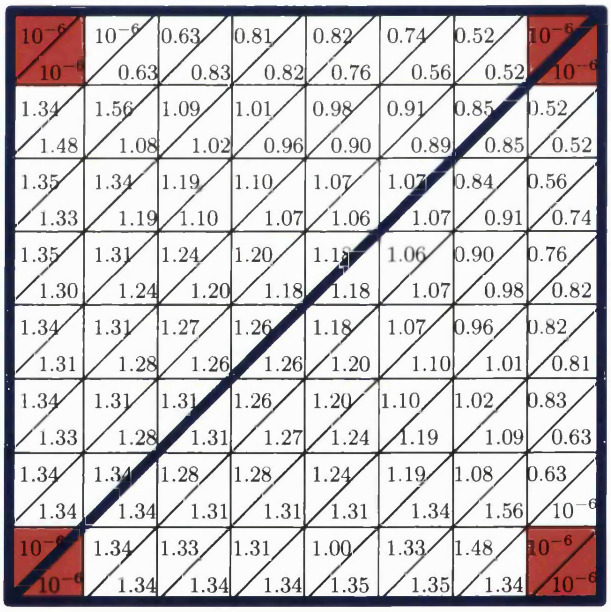
Figure 6-13: CG-iteration numbers for Experiment C for tolerance $\theta = 10^{-6}$, $h = 1/128$ and $\beta = 1/4$

A very similar behaviour can be seen, if we choose $\alpha = \hat{\alpha}$ in squares of size $2h \times 2h$ around the coarse grid nodes (see Figure 6-10). Computing the values $\left\|\nabla \psi_i^{Ms}|_\tau\right\|_2$ for all fine grid elements τ in two subdomains (as we had done in Figure 6-5 before) and choosing $\hat{\alpha} = 10^6$, we obtain Figure 6-14 (a), resp. 6-14 (b) for linear, resp. oscillatory boundary conditions for the construction of ψ_i^{Ms}. We can see that for the linear boundary conditions $\left\|\nabla \psi_i^{Ms}|_\tau\right\|_2 \approx H$ and therefore $\left\|\nabla \psi_i^{Ms}|_\tau\right\|_{2,\alpha} \approx 10^6 H$ on all of the red elements τ. For the oscillatory boundary conditions however, we get $\left\|\nabla \psi_i^{Ms}|_\tau\right\|_2 \approx 10^{-6} H$ and therefore $\left\|\nabla \psi_i^{Ms}|_\tau\right\|_{2,\alpha} \approx H$, i.e. the energy norm is kept constant in this case on the whole subdomain (this leads to well behaved $\gamma(\mathcal{V}_0, \alpha)$).

Similarly to the previous experiment if we choose the squares with $\alpha = \hat{\alpha}$ to be of size $4h \times 4h$, we cannot improve the convergence of the cg-method by the use of multiscale basis functions for $\beta H = h$ and $\beta H = 2h$, but find improvements for $\beta H = 3h$, squares of size $6h \times 6h$ the improvement can only be observed for $\beta H \geq 3h$, etc..



(a)



(b)

Figure 6-14: Values of $H^{-1} \cdot \left\| \nabla \psi_i^{Ms} |_\tau \right\|_2$ for Experiment C for linear boundary conditions (see case (a)), resp. for oscillatory boundary conditions (see case (b))

Experiment D:

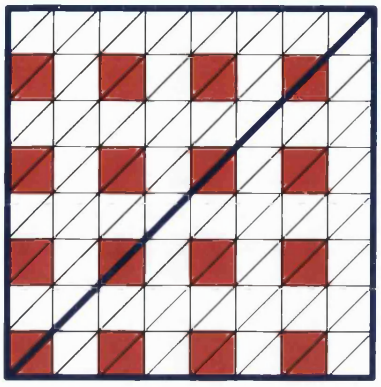


Figure 6-15: Subdomains for Experiment D (White elements: $\alpha^\tau = 1$, red elements: $\alpha^\tau = \hat{\alpha}$)

$\hat{\alpha}$	Lin. int.	Ms/Lin.b.c.	Ms/Osc.b.c.
10^{-11}	27	27	27
...	...	...	...
10^{-2}	27	27	27
10^{-1}	27	27	27
10^{0}	27	27	27
10^{1}	33	34	34
10^{2}	62	75	68
10^{3}	98	145	108
10^{5}	156	254	167
10^{7}	214	383	214
10^{9}	274	498	294

Figure 6-16: CG-iteration numbers for Experiment D for tolerance $\theta = 10^{-6}$, $h = 1/128$ and $\beta = 1/8$

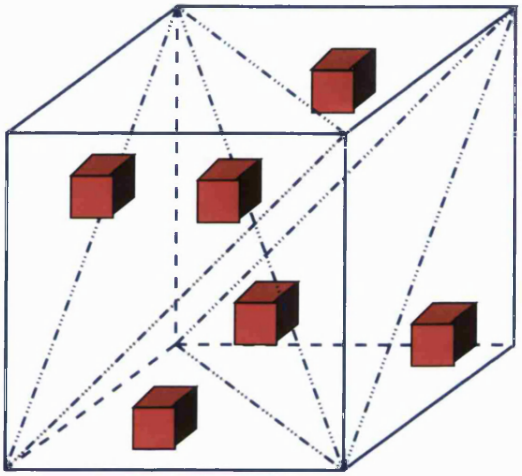
$\hat{\alpha}$	Lin. int.	Ms/Lin.b.c.	Ms/Osc.b.c.
10^{-11}	22	21	22
...	...	...	...
10^{-2}	22	21	22
10^{-1}	22	22	22
10^{0}	22	22	22
10^{1}	26	22	22
10^{2}	60	39	23
10^{3}	90	79	23
10^{5}	98	97	26
10^{7}	110	108	30
10^{9}	116	115	33

Figure 6-17: CG-iteration numbers for Experiment D for tolerance $\theta = 10^{-6}$, $h = 1/128$ and $\beta = 1/4$

Let us finally consider the case of many islands (see Figure 6-15) for $h = 1/128$, $H = 1/16$ and $\beta = 1/4$. Since in parts of the overlap we now have $\alpha = \hat{\alpha}$, Experiments B and C suggest we also might have to choose the overlap wisely. While $\beta H = h$ again leads to no or almost no improvement when using MsFE-interpolation (see Figure 6-16), we find a clear improvement with multiscale coarsening with oscillatory boundary conditions, when increasing this overlap (see Figure 6-17).

Example in 3D:

After these experiments for domains Ω in two dimensions, let us now consider three-dimensional problems. As a little model problem we choose a unit cube domain $\Omega = (0,1)^3$ with 24,576 fine grid elements, 48 subdomains, overlap $\beta H = 2h$ and one cube with $\alpha = \widehat{\alpha}$ per subdomain and $\alpha = 1$ on all other elements. As the linear and oscillatory boundary conditions are identical in this case, the cg-iteration numbers also are identical in both cases. Furthermore we see that while for the linear interpolation the cg-iteration numbers grow with $\hat{\alpha}$, they remain completely constant in the case of MsFE-interpolation (see Figure 6-18).



$\hat{\alpha}$	Lin. int.	Ms/Lin.b.c.
10^{-7}	25	25
...	...	...
10^{-1}	25	25
10^{0}	25	25
10^{1}	25	25
10^{2}	26	25
10^{3}	28	25
10^{5}	29	25
10^{7}	30	25
10^{9}	31	25
10^{11}	32	25
10^{13}	37	25

Figure 6-18: Convergence for 24,576 fine grid elements, 48 subdomains, overlap $\beta H = 2h$ and one cube with $\alpha = \widehat{\alpha}$ per subdomain and $\alpha = 1$ everywhere else (cube with 6 subdomains on the left and cg iteration numbers on the right)

An in-depth analysis of the convergence of domain-decomposition preconditioners with various coarsening techniques will be given in Graham et al. [60], which is currently in preparation.

6.3 Skeleton based MsFE-Interpolation

We now would like to find out what happens, if we extend the coarse grid space by introducing additional coarse grid basis functions. So far we always considered one coarse grid basis function for each coarse grid node $\mathbf{x}_i^H \in \mathcal{N}^H(\Omega)$. These were simple hat functions in the case of linear interpolation and multiscale basis functions with linear, resp. oscillatory boundary conditions in the case of MsFE-interpolation. Instead of using one coarse grid basis function for each coarse grid node, we could also use one for each skeleton node, i.e. for each node $\mathbf{x}_i^h$ of the fine grid that lies on any interior edge (resp. in 3D any interior edge or face) of the coarse grid (see Figure 6-19). This will increase the number of coarse grid basis functions to the number n_S of skeleton nodes significantly, will however remain (clearly) below the number of basis functions of the fine grid problem for H/h large enough. For 1D problems the skeleton consists only of the coarse grid nodes. The skeleton based method is therefore then identical for $d = 1$ to the (coarse grid) node based version as described earlier in this chapter.

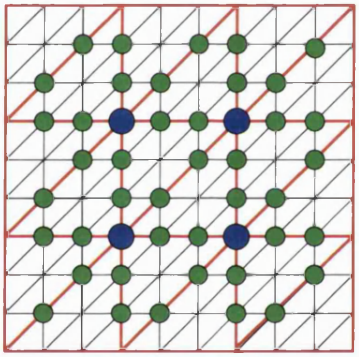


Figure 6-19: Notations: Fine grid (black), coarse grid (red), interior coarse grid nodes (blue), skeleton nodes (blue+green)

Let us now consider problems on a domain $\Omega \subset \mathbb{R}^d$, $d \geq 1$. Let Γ_j be the boundary of Ω_j and let $\mathbf{x}_i^h$ be a fine grid node on Γ_j. Then we could choose for the boundary conditions for the local skeleton based basis function $\vartheta_{i,\Omega_j}^{Skel}$

$$\vartheta_{i,\Omega_j}^{Skel}\left(\mathbf{x}_k^h\right) = \delta_{ik} \tag{6.85}$$

and interpolate linearly on the fine grid edges (resp. edges and faces) between these values on the boundary of Ω_j.

For each skeleton node $\mathbf{x}_i^h \in \Gamma_j$, the boundary of Ω_j, with the boundary conditions (6.85) we then solve as before

$$-\text{div}\left(\alpha(\mathbf{x})\,\nabla\vartheta_{i,\Omega_j}^{Skel}(\mathbf{x})\right) = 0 \text{ on } \Omega_j. \tag{6.86}$$

In a similar way as done before for the linear and node based multiscale basis functions, we can now define ψ_i^{Skel} globally by

$$\psi_i^{Skel}|_{\Omega_j} := \vartheta_{i,\Omega_j}^{Skel}, \tag{6.87}$$

The evaluation of the multiscale basis functions at the fine grid nodes then gives us the new coarse grid extension matrix

$$R_{0,Skel}^T := \left[\psi_1^{Skel}, ..., \psi_{n_S}^{Skel}\right] \tag{6.88}$$

and its transpose, the new coarse grid restriction matrix $R_{0,Skel}$.
This leads to a skeleton-based coarse grid stiffness matrix

$$A_{0,Skel} := R_{0,Skel} A R_{0,Skel}^T \tag{6.89}$$

and to

$$M_{2,Skel}^{-1} := \sum_{i=1}^{p} R_i^T A_i^{-1} R_i + R_{0,Skel}^T A_{0,Skel}^{-1} R_{0,Skel} = M_1^{-1} + R_{0,Skel}^T A_{0,Skel}^{-1} R_{0,Skel}. \tag{6.90}$$

This new approach does not lead to a new preconditioner for $A(\alpha)$, but (6.90) turns out to be another way of writing the inverse of $A(\alpha)$ if $\widetilde{\Omega_j} = \Omega_j$ for all $j = 1, ..., p$. This therefore can be very useful if we would like to parallelise exact solvers for the groundwater flow problem.

THEOREM 6.6
If $\widetilde{\Omega_j} = \Omega_j$ for all $j = 1, ..., p$ (i.e. no overlap) then the two-level additive Schwarz preconditioner $M_{2,Skel}^{-1}$ with multiscale coarsening in one dimension and the two-level skeleton-based additive Schwarz preconditioner with multiscale coarsening in higher dimensions are exact inverses of the stiffness matrix $A = A(\alpha)$.

Proof:
Let $\mathcal{S}$ denote the set of indices of points which lie on the skeleton inside the domain Ω (which is identical to the coarse grid nodes $\mathcal{N}^H(\Omega)$ for one-dimensional domains) and let $\mathcal{I}$ denote the set of indices of points inside the domain, which do not lie on the

skeleton and let n_S be the number of skeleton nodes in Ω. Then we can reorder the rows and columns of A such that A_{SS} is the matrix block of contributions of freedoms on the skeleton, etc., to get

$$A = \begin{bmatrix} A_{\mathcal{I}\mathcal{I}} & A_{\mathcal{I}S} \\ A_{S\mathcal{I}} & A_{SS} \end{bmatrix}. \tag{6.91}$$

Now let $\bar{\Omega}_j$ be an arbitrary coarse grid element. Furthermore let $\mathcal{I}_j$ denote the index set corresponding to the nodes in the interior of Ω_j. Then $I = \bigcup_{j=1}^{p} I_j$ and $A_{\mathcal{I}\mathcal{I}}$ is block diagonal with a diagonal block $A_{\mathcal{I}_j\mathcal{I}_j}$ for each Ω_j.
Furthermore let ψ_l^{Skel} be the l-th multiscale basis function on Ω and $\boldsymbol{\psi}_l^{Skel}$ the vector of its evaluations at the non-skeleton nodes in Ω. Then the values of $\boldsymbol{\psi}_l^{Skel}$ are given by the solution of the equation

$$(A_{\mathcal{I}\mathcal{I}})\, \boldsymbol{\psi}_l^{Skel} = -\,(A_{\mathcal{I}S})\, \mathbf{e}_l, \tag{6.92}$$

where $\mathbf{e}_l$ is the l-th standard basis vector in $\mathbb{R}^{n_S}$.
(Note that this implies that ψ_l^{Skel} is zero on any coarse grid element, which does not contain the l-th skeleton node.)

Writing this result in matrix form, we get

$$\left[\boldsymbol{\psi}_1^{Skel}, \boldsymbol{\psi}_2^{Skel}, ..., \boldsymbol{\psi}_{n_S}^{Skel}\right] = -\,(A_{\mathcal{I}\mathcal{I}})^{-1}\,(A_{\mathcal{I}S})\,. \tag{6.93}$$

Let us now consider the operation of multiscale interpolation from the coarse grid to the fine grid. By (6.93) and by definition (6.88) of the coarse grid extension matrix $R_{0,Skel}^T$, we obtain immediately

$$R_{0,Skel}^T = \begin{bmatrix} -A_{\mathcal{I}\mathcal{I}}^{-1} A_{\mathcal{I}S} \\ \mathbb{I} \end{bmatrix}, \tag{6.94}$$

where $\mathbb{I}$ denotes an identity matrix, indicating that values at skeleton nodes are left unchanged by the extension matrix.

Writing A in block form as in (6.91), we then get

$$A_{0,Skel} := R_{0,Skel} A R_{0,Skel}^T = A_{SS} - A_{S\mathcal{I}} A_{\mathcal{I}\mathcal{I}}^{-1} A_{\mathcal{I}S} =: S, \tag{6.95}$$

where S is called the Schur complement of $A_{\mathcal{SS}}$, and

$$\begin{aligned} R_{0,Skel}^T A_{0,Skel}^{-1} R_{0,Skel} &= \begin{bmatrix} -A_{\mathcal{II}}^{-1} A_{\mathcal{IS}} \\ \mathbb{I} \end{bmatrix} \overbrace{\left[A_{\mathcal{SS}} - A_{\mathcal{SI}} A_{\mathcal{II}}^{-1} A_{\mathcal{IS}}\right]^{-1}}^{S^{-1}} \begin{bmatrix} -\left(A_{\mathcal{II}}^{-1} A_{\mathcal{IS}}\right)^T & \mathbb{I} \end{bmatrix} \\ &= \begin{bmatrix} A_{\mathcal{II}}^{-1} A_{\mathcal{IS}} S^{-1} A_{\mathcal{SI}} A_{\mathcal{II}}^{-1} & -A_{\mathcal{II}}^{-1} A_{\mathcal{IS}} S^{-1} \\ -S^{-1} A_{\mathcal{SI}} A_{\mathcal{II}}^{-1} & S^{-1} \end{bmatrix}. \end{aligned} \tag{6.96}$$

In the case of non-overlapping domains Ω_j the two-level skeleton-based additive Schwarz preconditioner with multiscale interpolation then is

$$\begin{aligned} M_{2,Skel}^{-1} &= \begin{bmatrix} A_{\mathcal{II}}^{-1} & 0 \\ 0 & 0 \end{bmatrix} + R_{0,Skel}^T A_{0,Skel}^{-1} R_{0,Skel} \\ &= \begin{bmatrix} A_{\mathcal{II}}^{-1} + A_{\mathcal{II}}^{-1} A_{\mathcal{IS}} S^{-1} A_{\mathcal{SI}} A_{\mathcal{II}}^{-1} & -A_{\mathcal{II}}^{-1} A_{\mathcal{IS}} S^{-1} \\ -S^{-1} A_{\mathcal{SI}} A_{\mathcal{II}}^{-1} & S^{-1} \end{bmatrix}. \end{aligned} \tag{6.97}$$

Therefore

$$\begin{aligned} M_{2,Skel}^{-1} A &= \begin{bmatrix} A_{\mathcal{II}}^{-1} + A_{\mathcal{II}}^{-1} A_{\mathcal{IS}} S^{-1} A_{\mathcal{SI}} A_{\mathcal{II}}^{-1} & -A_{\mathcal{II}}^{-1} A_{\mathcal{IS}} S^{-1} \\ -S^{-1} A_{\mathcal{SI}} A_{\mathcal{II}}^{-1} & S^{-1} \end{bmatrix} \begin{bmatrix} A_{\mathcal{II}} & A_{\mathcal{IS}} \\ A_{\mathcal{SI}} & A_{\mathcal{SS}} \end{bmatrix} \\ &= \mathbb{I}, \end{aligned} \tag{6.98}$$

i.e. $M_{2,Skel}^{-1}$ is then the exact inverse of A. □

If the overlap is chosen greater than 0, we will not obtain the exact inverse anymore and will increase the computational costs. This can also be seen in the numerical results in Subsection 7.2.2.

6.4 Summary

In this chapter we introduced a new coarsening operator for the two-level additive Schwarz preconditioner. It is based on multiscale finite elements and therefore contains information of the underlying permeability field. We discussed possible boundary conditions for MsFE basis functions for two- and three-dimensional problems and implementation aspects of the new operator.

Furthermore we analysed the convergence properties of additive Schwarz preconditioners with MsFE coarsening operators for a model problem in which the coefficient α was supposed to be constant within the overlap of subdomains. This analysis was based on the so called coarse space robustness indicator (depending on the coarse grid basis, the coarse mesh and α), which (for the model problem) is small, when we observe fast convergence, and large, when the preconditioned conjugate gradient method converges slowly. After this strict analysis we then considered more general model problems in which next to the coarsening operator also the size of the overlap between the subdomains has to be chosen in the right way (depending on α) to achieve an optimal rate of convergence when applying a two-level additive Schwarz preconditioner to their system matrix in the cg-method. Numerical results in 2D and 3D for two media problems were given, that support the theory.
We finally enriched the coarse grid space by introducing a basis function for every node of the skeleton. It was shown that the new additive Schwarz preconditioner is the exact inverse of the stiffness matrix $A(\alpha)$.
The new multiscale coarsening operators with coarse grid and skeleton based interpolation will be compared with linear coarsening in quite some detail in the next chapter for random field problems. We will give numerical results comparing cg-iteration numbers as well as computation times. Furthermore we will give an analysis of the complexity of the setup and use of these preconditioners in a cg-method.

Chapter 7

Efficiency and Applications

"*Efficiency is intelligent laziness.*"

— David Dunham (Software developer and entrepreneur) —

"*I have been impressed with the urgency of doing. Knowing is not enough; we must apply. Being willing is not enough; we must do.*"

— Leonardo da Vinci quotes (Italian draftsman, painter, sculptor, architect and engineer, 1452-1519) —

"*An ounce of performance is worth pounds of promises.*"

— Mae West (American actress and sex symbol, 1892-1980) —

7.1 Short overview of the theoretical results

In Chapter 4 we derived upper bounds for the condition numbers of the stiffness matrix $A(\alpha)$, obtained from PDE problem (2.13,2.14). In Theorem 4.1 we showed:

$$\kappa\left(A(\alpha)\right) \leq \max_{\mathbf{x},\mathbf{y}\in\Omega} \left(\frac{\alpha\left(\mathbf{x}\right)}{\alpha\left(\mathbf{y}\right)}\right) \kappa\left(A(1)\right). \tag{7.1}$$

This possible dependence on the global ratio of α was then reduced in Chapter 5 by focusing on additive Schwarz preconditioners $M_{2,Lin}^{-1}$ with linear coarsening. For the preconditioned system $M_{2,Lin}^{-1}A(\alpha)$ we proved the following bound in Theorem 5.16

$$\kappa\left(M_{2,Lin}^{-1}A(\alpha)\right) \leq C\left(1+\beta^{-1}\right)\left(\max_{j}\max_{\mathbf{x},\mathbf{y}\in\Omega_j}\left(\frac{\alpha(\mathbf{x})}{\alpha(\mathbf{y})}\right)\right)B(d), \tag{7.2}$$

where $B(1) := 1$, $B(2) := (1+\log(H/h))$ and $B(3) := (H/h)$ and $0 < \beta < 1$.
In Theorem 5.17 we showed

$$\kappa\left(M_{2,Lin}^{-1}A(\alpha)\right) \leq C\left(1+\beta^{-1}\right)\left(\max_{j}\max_{\mathbf{x},\mathbf{y}\in\omega_{\Omega_j}}\left(\frac{\alpha(\mathbf{x})}{\alpha(\mathbf{y})}\right)\right). \tag{7.3}$$

While bound (7.2) is better than (7.3) with respect to the permeability α, bound (7.3) gets rid of the dependence on the mesh ratio H/h. This can be especially important for $3D$ problems, because of the definition of $B(3)$.
To reduce the dependence on α even further we introduced a new coarsening operator in Chapter 6, that we based on the concept of multiscale finite elements. By taking the permeability field into consideration for the coarse grid problem in this way, we managed to completely remove the dependence on α, if the permeability is constant inside all overlaps. For more complex fields, the convergence also depends among others on the choice of boundary conditions for the construction of the MsFE-basis functions and on the overlap βH. While for the case of constant α inside the overlap we gave theoretical bounds for the condition number of the preconditioned system, $M_{2,Ms}^{-1}A(\alpha)$, the analysis gets more complex for general fields.
After the discussion of these preconditioners, we proved in Theorem 6.6 that enriching the coarse space by the introduction of additional basis functions for each skeleton node leads to an exact inverse $M_{2,Skel}^{-1}$ of $A(\alpha)$ (for zero overlap). While this is computationally as expensive as computing the exact inverse in an other way (for example by Gaussian elimination), it does however provide us with an interesting new parallelisation strategy by splitting the problem into many local solves.
Between the node based preconditioners and the skeleton based version there is still some space for preconditioners, which by inputting some extra work compared to node based preconditioners could bring an improvement in terms of the dependence on h, H and α compared to them, but which still comes with a reduced setup cost and cost per iteration compared to the skeleton based version.
These immediate preconditioners might for example be based on preconditioning the Schur complement S of $A_{\mathcal{SS}}$ (as defined in Section 6.3). This could well be part of an interesting future project that might come out of this thesis.

7.2 Numerical results for random permeability fields

7.2.1 One dimensional problems

Let us now check the convergence of the preconditioned conjugate gradient method by considering some numerical results. Let us therefore start by checking for random fields of Ornstein-Uhlenbeck type on $\Omega = (0,1)$ the condition numbers of the two level additive Schwarz preconditioned systems with linear, resp. (node-based and skeleton-based) multiscale interpolation, as well as the cg-iteration numbers, when using these preconditioners. Fix therefore $h = 1/256$, resp. $h = 1/512$, $\lambda = 10h$ and $\beta H = h$.

Degrees of freedom	Variance σ^2	Linear interpolation $\kappa\left(M_{2,Lin}^{-1}A(\alpha)\right)$	Its.	Multiscale interpolation $\kappa\left(M_{2,Ms}^{-1}A(\alpha)\right)$	Its.
255	1	$1.9 \cdot 10^3$	8	$1.1 \cdot 10^1$	3
	2	$3.9 \cdot 10^3$	9	$1.1 \cdot 10^1$	3
	4	$8.8 \cdot 10^3$	10	$1.1 \cdot 10^1$	3
	6	$2.7 \cdot 10^4$	13	$1.1 \cdot 10^1$	3
	16	$1.1 \cdot 10^5$	19	$1.2 \cdot 10^1$	4
	32	$5.4 \cdot 10^5$	44	$1.2 \cdot 10^1$	4
511	1	$7.4 \cdot 10^3$	8	$1.1 \cdot 10^1$	3
	2	$1.6 \cdot 10^4$	9	$1.1 \cdot 10^1$	3
	4	$3.9 \cdot 10^4$	10	$1.1 \cdot 10^1$	3
	8	$1.3 \cdot 10^5$	15	$1.2 \cdot 10^1$	3
	16	$5.6 \cdot 10^5$	25	$1.2 \cdot 10^1$	4
	32	$2.9 \cdot 10^6$	53	$1.2 \cdot 10^1$	4

Table 7.1: Average condition numbers and average cg-iteration numbers of the preconditioned systems for 1,000 fields with correlation of Ornstein-Uhlenbeck type with correlation length $\lambda = 10h$ for overlap $\beta H = h$

We compute the condition numbers and numbers of iterations for 1000 fields and give the average values in table 7.1. We observe that for linear interpolation both $\kappa\left(M_{2,Lin}^{-1}A(\alpha)\right)$ and the number of cg-iterations grow strongly with σ^2 as expected by the theory. For MsFE-interpolation (i.e. here $M_{2,Ms}^{-1} = M_{2,Skel}^{-1}$) on the other hand both stay almost constant and close to 1. (We obtain values slightly bigger than 1 here since we consider overlapping subdomains and the MsFE-preconditioner therefore is not the exact inverse of the global stiffness matrix A.)

7.2.2 Two dimensional problems

h	Var.	$\lambda = 1/8$			$\lambda = 1/16$		
	σ^2	Lin. Int.	Ms. Int.	Skel. b.	Lin. Int.	Ms. Int.	Skel. b.
1/64	1	25	23	3	26	24	3
	2	27	25	3	30	27	3
	4	31	27	3	36	30	3
	6	34	30	3	43	34	3
	8	38	32	3	50	39	3
	12	46	36	3	66	47	3
	16	54	40	3	88	57	3
	20	63	45	3	111	68	3
1/128	1	28	25	4	29	27	4
	2	30	27	4	35	30	4
	4	37	31	4	47	36	4
	6	43	34	4	62	41	4
	8	50	37	4	79	47	4
	12	66	44	4	121	61	4
	16	86	52	4	178	77	4
	20	108	60	4	247	98	4
1/256	1	30	27	4	32	29	4
	2	35	29	4	40	33	4
	4	47	34	4	59	42	4
	6	63	39	4	84	52	4
	8	83	48	4	117	62	4
	12	133	64	4	209	85	4
	16	195	81	4	322	110	4
	20	282	102	4	451	139	4

Table 7.2: Average cg-iteration numbers for 100 fields, $H = 1/16$ and overlap $\beta H = 1/64$

For two-dimensional problems with domain $\Omega = (0,1)^2$, we expect to obtain exact inverses in the case of skeleton based interpolation and no overlap (i.e. $\beta = 0$). Let us now consider the cg-iteration numbers for $h = 1/64$, $h = 1/128$, resp. $h = 1/256$, $H = 1/16$ and Overlap $\beta H = 1/64$ for 100 random fields with correlation lengths which are smaller, equal, resp. larger than the size of the subdomains (see Figures 7.2 and 7.3). As expected by Theorem 6.6, the additive Schwarz preconditioners with skeleton based MsFE-interpolation $M^{-1}_{2,Skel}$ lead to almost immediate convergence. (However the iteration numbers are greater than 1, since the overlap in these experiments was chosen greater than 0.) For node based MsFE-interpolation (with oscillatory boundary

conditions for the construction of the coarse grid basis functions) the average number of cg-iterations is significantly below those obtained for linear interpolation especially for large variance σ^2 (up to a factor 5 smaller for our experiments).

h	Var.	$\lambda = 1/32$			$\lambda = 1/64$		
	σ^2	Lin. Int.	Ms. Int.	Skel. b.	Lin. Int.	Ms. Int.	Skel. b.
1/64	1	29	27	3	31	28	3
	2	34	30	3	39	34	3
	4	47	37	3	58	44	3
	6	61	43	3	85	57	3
	8	76	50	3	115	73	3
	12	116	62	3	196	96	3
	16	165	78	3	319	123	3
	20	234	107	3	469	172	3
1/128	1	31	29	4	32	30	4
	2	40	33	4	41	36	4
	4	59	42	4	63	48	4
	6	85	52	4	95	65	4
	8	118	62	4	140	80	4
	12	210	86	4	269	112	4
	16	340	114	4	463	146	4
	20	511	147	4	763	201	4
1/256	1	32	30	4	33	31	4
	2	43	34	4	45	37	4
	4	64	43	4	70	52	4
	6	94	54	4	103	72	4
	8	141	65	4	163	88	4
	12	259	91	4	312	130	5
	16	512	124	5	722	170	5
	20	822	172	5	$> 1,000$	237	5

Table 7.3: Average cg-iteration numbers for 100 fields, $H = 1/16$ and overlap $\beta H = 1/64$

7.2.3 Three dimensional problems

The same sort of behaviour can also be observed for three dimensional domains $\Omega = (0,1)^3$. Here we consider problems with 24,576 fine grid elements and two sets of parameters H and β, first $H = 8h$ and $\beta H = 2h$ and then $H = 4h$ and $\beta H = h$ in both cases with linear boundary conditions for the construction of the MsFE coarse grid basis functions. In both cases we see a reduction of the cg-iteration numbers even for

this small model problem when using MsFE-interpolation. Using oscillatory boundary conditions for the MsFE basis functions, these iteration numbers are likely to reduce further.

Var. σ^2	Lin. int.	Ms/Lin.b.c.
1	27	27
2	28	27
4	29	28
6	31	29
8	32	30
16	36	33
32	45	38

Var. σ^2	Lin. int.	Ms/Lin.b.c.
1	41	38
2	49	40
4	72	61
6	100	79
8	131	99
16	308	229
32	869	509

Figure 7-1: CG iteration numbers for 24,576 fine grid elements and for 48 subdomains and overlap $\beta H = 2h$ (on the left) and 384 subdomains and overlap $\beta H = h$ (on the right) for completely random three dimensional random fields

7.3 Complexity

In this subsection we shall estimate theoretically how much computation will be done in the cg-iterations as well as when setting up the different two-level additive Schwarz preconditioners, that we defined earlier in the thesis. This analysis will be reflected in the actual computation times reported in Subsection 7.4. The *complexity analysis* can be performed in several different ways. In this thesis we limit our analysis to counting the arithmetic operations and will for simplicity ignore all time needed for memory access.

The overall complexity of a (unpreconditioned) conjugate gradient method is well known. It is therefore more interesting for us to find out how much more the use of MsFE-interpolation instead of linear interpolation in a two-level additive Schwarz method costs, and whether the MsFE-interpolation increases the order of the complexity of the setup of the matrices involved in the computations or not.

For simplicity consider a domain $\Omega = (0,1)^d$ with fine grid nodes $(i_1/n, i_2/n, ..., i_d/n)$, $i_1, i_2, ..., i_d \in \{0, ..., n\}$, and coarse grid nodes $(j_1/m, j_2/m, ..., j_d/m)$, $j_1, j_2, ..., j_d \in \{0, ..., m\}$, $m < n$, where n/m is an integer, which are the nodes of a uniform fine and coarse grid triangulation (i.e. intervals in 1D, triangles in 2D and tetrahedra in 3D). Let $f(n) = \mathcal{O}(g(n))$, if $\lim_{n\to\infty} |f(n)/g(n)| < \infty$. We are then interested in the complexity bounds when $m, n \to \infty$.

With this notation the domain contains $\mathcal{O}\left(n^d\right)$ fine and $\mathcal{O}\left(m^d\right)$ coarse grid elements.

For the following complexity analysis, we would like to consider the following costs:

- **Setup cost:**
  - Cost of finding the local (overlapping subdomain) matrices
  - Cost of finding the coarse grid matrix in the case of
    * linear interpolation
    * MsFE-interpolation
  - Cost of factorising the local (overlapping subdomain) matrices
  - Cost of factorising the coarse grid stiffness matrix
- **Iterative solution cost:**
  - Cost of one cg iteration (requiring backsolves with the factorised local and coarse grid matrices).

(A) Matrix sizes:
Let us for the moment concentrate on a subdomain $\widetilde{\Omega_j}$ having an overlap βH with its neighbouring subdomains, where $0 < \beta < 1$. In this case the number of fine grid nodes in $\widetilde{\Omega_j}$ is therefore $\mathcal{O}\left((n/m)^d\right)$. The coarse grid has got $\mathcal{O}\left(m^d\right)$ nodes.

(B) Setup cost:

- **Setup of the local matrices:**
  The local restriction matrices R_j and extension matrices R_j^T do not depend on the sort of coarsening used in the preconditioner. Let n_j denote the number of inner nodes in $\widetilde{\Omega_j}$, then R_j^T has got $\mathcal{O}\left(n^d \times n_j\right)$ entries. An entry $\left(R_j^T\right)_{kl}$ of R_j^T is 1 if and only if $x_k^h \in \widetilde{\Omega_j}$ and 0 otherwise. This means that for the setup of the local extension and restriction matrices and therefore also of the local stiffness matrices no computations are necessary. The matrices R_j and R_j^T, $j = 1, ..., p$, in fact do not have to be set up explicitly (since we only have to solve linear systems with matrices A_j, $j = 0, ..., p$). Since we ignore memory access setting up the matrices A_j, $j = 1, ..., p$, with $A_j := R_j A R_j^T$ is therefore also of zero complexity.

- **Setup of the global matrices:**
  We next shall compare the setup cost of the coarse grid stiffness matrices A_0 (linear interpolation) and $A_{0,Ms}$ (MsFE interpolation). Let A^0_{kl}, resp. $A^{0,Ms}_{kl}$ be the entries of A_0, resp. of $A_{0,Ms}$. They are given by

$$\begin{aligned} A^0_{kl} &:= \int_\Omega \alpha \nabla \psi^{Lin}_k . \nabla \psi^{Lin}_l \\ &= \sum_{\tau \subset \mathrm{supp}(\psi^{Lin}_k) \cap \mathrm{supp}(\psi^{Lin}_l)} \int_\tau \alpha_\tau \nabla \psi^{Lin}_k . \nabla \psi^{Lin}_l \\ &= \sum_{\tau \subset \mathrm{supp}(\psi^{Lin}_k) \cap \mathrm{supp}(\psi^{Lin}_l)} |\tau| \, \alpha_\tau \nabla \psi^{Lin}_k (\mathbf{m}^\tau) \cdot \nabla \psi^{Lin}_l (\mathbf{m}^\tau) , \end{aligned} \tag{7.4}$$

  assuming α is constant and equal to α^τ on each τ and where $\mathbf{m}^\tau$ is the centroid of fine grid element τ, resp. by

$$A^{0,Ms}_{kl} := \sum_{\tau \subset \mathrm{supp}(\psi^{Ms}_k) \cap \mathrm{supp}(\psi^{Ms}_l)} |\tau| \, \alpha_\tau \nabla \psi^{Ms}_k (\mathbf{m}^\tau) \cdot \nabla \psi^{Ms}_l (\mathbf{m}^\tau) \tag{7.5}$$

  (since $\nabla \psi^{Lin}_k$, $\nabla \psi^{Lin}_l$, $\nabla \psi^{Ms}_k$ and $\nabla \psi^{Ms}_l$ are constant on each fine grid element). As there are $\mathcal{O}((n/m)^d)$ fine grid elements in $\mathrm{supp}\left(\psi^{Ms}_k\right) \cap \mathrm{supp}\left(\psi^{Ms}_l\right)$, the setup cost of each non-zero value A^0_{kl}, resp. $A^{0,Ms}_{kl}$ (assuming that the values of the linear, resp. multiscale coarse grid functions at the fine grid nodes are given) is $\mathcal{O}((n/m)^d)$ for linear and multiscale interpolation due to the sum over the fine grid elements in (7.4), resp. (7.5). (Here the multiplicative factor of the cost, which is hidden in the $\mathcal{O}$-notation can be larger in the multiscale case, since $\nabla \psi^{Lin}_k$ and $\nabla \psi^{Lin}_l$ are constant on each subdomain, while $\nabla \psi^{Ms}_k$ and $\nabla \psi^{Ms}_l$, which depend on α, will usually only be constant on the fine grid elements.)

But what are the additional costs for *setting up the coarse grid basis functions*? The linear coarse grid basis functions ψ^{Lin}_i, $\mathbf{x}^H_i \in \mathcal{N}^H(\Omega)$, (resp. the vector of their evaluations at the fine grid nodes) only have to be computed once and can then be used for all underlying fields. As in practice we have to solve the problem for many fields to then be able to apply some sort of Monte Carlo method, we can ignore the cost of evaluating ψ^{Lin}_k at the fine grid nodes and their setup cost and can assume that they are known. The cost of computing a non-zero entry A^0_{kl} is therefore $\mathcal{O}\left((n/m)^d\right)$. Let b be the maximum number of fine grid nodes, that share a common edge with a fine grid node in the interior of Ω. Since b (independent of n and m) is then also an upper bound for the number of coarse grid nodes $\mathbf{x}^H_l$, that share a common edge with a given coarse grid node $\mathbf{x}^H_k$, the

number of non-zero entries of A_0, resp. $A_{0,Ms}$ is bounded by $\mathcal{O}\left(m^d\right)$. This finally leads to the setup cost of A_0 of $\mathcal{O}\left(n^d\right)$.

In the case of MsFE-interpolation the computation of the coarse grid basis functions ψ_k^{Ms} is more expensive. Their setup cost cannot be neglected, as ψ_k^{Ms} depends on the underlying field and has to be computed for each field separately. On a subdomain Ω_i we have to find D basis functions, where $D \leq d+1$ and where $D = d+1$ for the subdomains that do not touch the boundary Γ and $D < d+1$ for subdomains touching Γ assuming zero boundary conditions, i.e. $u(\mathbf{x}) = 0$, $\forall \mathbf{x} \in \Gamma$. We can find each of them by solving a local linear problem.
We therefore have to set up a local stiffness matrix of size $\mathcal{O}\left((n/m)^d \times (n/m)^d\right)$ for each subdomain, for which the setup cost (see the bound for the local stiffness matrices A_i) is of complexity 0. Furthermore, we have to set up a different right-hand-side vector for each of the D basis functions. These vectors depend on the boundary conditions that we choose for the construction of the MsFE basis functions. For 2D problems we gave explicit formulae for oscillatory boundary conditions (see subsection 6.13). The setup cost for the right-hand-side vectors does therefore only vary slightly compared to linear boundary conditions. For $d > 2$ we have to solve $(d-1)$-dimensional subproblems for the construction of the boundary conditions. This can be expensive, is however of cost of lower order w.r.t. n and m compared to the computation cost of the values of the MsFE basis functions in the interior of the subdomains. We therefore can in all cases ignore the setup cost of these right-hand-side vectors.

One occurring problem now is the fact, that if we choose small subdomains, we only have to solve small local problems, but the coarse grid problem then is large. On the other hand if we keep the coarse grid problem small, we have to solve larger local problems. Let us however assume that we can choose m in such a way, that the local and coarse grid stiffness matrices are small enough compared to the global stiffness matrix, so that we can solve the local and the coarse grid systems directly. When we have to solve a linear system with a $r \times r$ symmetric positive definite system matrix $\mathcal{A}$ for some $r \in \mathbb{N}$ we could use for example the so called *Cholesky-decomposition*, i.e. we factorize

$$\mathcal{A} = LL^T, \tag{7.6}$$

where L is a lower triangular matrix. The cost of computing the factorization

of $\mathcal{A}$ in the Cholesky decomposition is $\mathcal{O}(r^3)$ for $r \to \infty$. When applying this decomposition to solve a linear system, we have to perform two backsolves which have complexity $\mathcal{O}(r^2)$ for $r \to \infty$.

For a sparse matrix $\mathcal{A}$ (as in our case) this computational cost can however be reduced significantly (depending on the sparsity pattern of $\mathcal{A}$). One direct method to use is a *banded Cholesky solver* (which is suboptimal for our problems, but gives some first idea of possible improvements, see Demmel [37], Chapter 6). Its complexity depends on the *bandwidth* B of the matrix, i.e. the maximum number of subdiagonals resp. supdiagonals that non-zero entries are away from the main diagonal. Let us only consider the case of two dimensional domains as an example here (similar bounds can of course also be found for other dimensions). Computing a factorization of the system matrix using this direct solver costs $\mathcal{O}\left(rB^2\right)$, where for our matrices for $d = 2$ we have $B = \sqrt{r}$. Therefore using a banded solver reduces the factorization cost for $d = 2$ from $\mathcal{O}\left(r^3\right)$ of the Cholesky decomposition to $\mathcal{O}\left(r^2\right)$ for $r \to \infty$. The cost for one backsolve then reduces from $\mathcal{O}(r^2)$ to $\mathcal{O}(rB) = \mathcal{O}(r^{3/2})$.
However this is still not optimal yet. Many different methods have been developed in the last few decades and we do not want to compare these methods in this thesis or express our preference for any one of them. We therefore denote $\mathcal{S}(r)$ for the cost of the factorising of a $r \times r$ sparse system matrix and $\mathcal{B}(r)$ for the cost of one backsolve using this factorisation. The sparsity patterns of all stiffness matrices considered in this subsection are similar, as they are based on the same uniform fine grid and coarse grid and the same second order differential operator (with different coefficients α perhaps). A complexity analysis without going into any detail of the exact type of direct solver therefore seems reasonable.

As for the computation of the multiscale basis functions we are dealing with sparse (local) matrices of size $\mathcal{O}\left((n/m)^d \times (n/m)^d\right)$ for the computation of the $\mathcal{O}\left((n/m)^d\right)$ values of ψ_k^{Ms} in the interior of Ω_j, the cost of solving each of these sparse systems is then given by $\mathcal{O}\left(\mathcal{S}\left((n/m)^d\right)\right)$. As Ω consists of m^d subdomains, the cost to compute the basis functions is therefore then of order $\mathcal{O}\left(m^d \cdot \mathcal{S}\left((n/m)^d\right)\right)$.
Setting up the matrices for the additive Schwarz preconditioner with multiscale interpolation is hence of order $\mathcal{O}\left((n/m)^{2d} + m^d \cdot \mathcal{S}\left((n/m)^d\right)\right)$.
This shows that the extra cost of setting up MsFE-coarse grid matrices instead of linear interpolation matrices is given by $\mathcal{O}\left(m^d \cdot \mathcal{S}\left((n/m)^d\right)\right)$ for general sparse

direct solvers and by $\mathcal{O}\left(m^{-2}n^4\right)$ for a banded solver for $d = 2$.

- **Cost of factorising the local (overlapping subdomain) matrices**
  Having set up the matrices, we now have to factorise $\mathcal{O}\left(m^d\right)$ local stiffness matrices (one for each subdomain $\widetilde{\Omega_j}$ with (sparse) system matrices of size $\mathcal{O}\left((n/m)^d \times (n/m)^d\right)$ and one problem on the coarse grid with (sparse) system matrix of size $\mathcal{O}\left(m^d \times m^d\right)$. As we have to factorise $\mathcal{O}\left(m^d\right)$ local stiffness matrices, applying the local part of the preconditioner therefore has got a cost of $\mathcal{O}\left(m^d \cdot \mathcal{S}\left((n/m)^d\right)\right)$ for general direct solvers, resp. of $\mathcal{O}\left(m^{-2}n^4\right)$ for banded Cholesky solvers and $d = 2$.

- **Cost of factorising the coarse grid stiffness matrix**
  Applying the coarse part costs $\mathcal{O}\left(\mathcal{S}\left(m^d\right)\right)$, resp. $\mathcal{O}\left(m^4\right)$, which means that the overall cost for factorising stiffness matrices is given by $\mathcal{O}\left(m^d \cdot \mathcal{S}\left((n/m)^d\right) + \mathcal{S}\left(m^d\right)\right)$ for general direct solvers, resp. by $\mathcal{O}\left(m^{-2}n^2 + m^4\right)$ for banded Cholesky and $d = 2$.

(C) Cost of applying the preconditioner in the cg-method:
We now would like to compute the cost of applying the different preconditioners in the following version of the *preconditioned conjugate gradient method*:

(i) *Choose an initial guess* $\mathbf{x}_0$.

(ii) *Compute* $\mathbf{r}_0 := \mathbf{b} - A\mathbf{x}_0$, $\mathbf{z}_0 := M^{-1}\mathbf{r}_0$ *and* $\mathbf{p}_0 := \mathbf{z}_0$.

(iii) ***For*** $k = 0, 1, \ldots,$ ***until*** *convergence* ***do:***

(iv) $\alpha_k := (\mathbf{r}_k, \mathbf{z}_k) \,/\, (A\mathbf{p}_k, \mathbf{p}_k)$

(v) $\mathbf{x}_{k+1} := \mathbf{x}_k + \alpha_k \mathbf{p}_k$

(vi) $\mathbf{r}_{k+1} := \mathbf{r}_k - \alpha_k A\mathbf{p}_k$

(vii) $\mathbf{z}_{k+1} := M^{-1}\mathbf{r}_{k+1}$

(viii) $\beta_k := (\mathbf{r}_{k+1}, \mathbf{z}_{k+1}) \,/\, (\mathbf{r}_k, \mathbf{z}_k)$

(ix) $\mathbf{p}_{k+1} := \mathbf{z}_{k+1} + \beta_k \mathbf{p}_k$

(x) ***end do.***

In the following analysis, $M^{-1} := M_{2,Lin}^{-1}$ in the case of linear interpolation and $M^{-1} := M_{2,Ms}^{-1}$ for multiscale interpolation.

The dominating part in terms of complexity in a cg-iteration iteration is step (vii), i.e. the backsolves with preconditioner M^{-1}. Let us assume that the local matrices, resp. coarse grid matrix, that form the preconditioner were already factorised in the setup of the M^{-1}. We therefore only have to perform the (cheaper) backsolves instead of the whole factorising inside the cg-method.
As we have to solve $\mathcal{O}\left(m^d\right)$ local problems, the backsolves in the case of applying the local part of the preconditioner therefore has got a cost of $\mathcal{O}\left(m^d \cdot \mathcal{B}\left((n/m)^d\right)\right)$ for general direct solves, resp. of $\mathcal{O}\left(m^2 \cdot (n/m)^3\right) = \mathcal{O}\left(m^{-1} \cdot n^3\right)$ for banded Cholesky solvers for $d = 2$. Furthermore we have to solve one coarse grid problem which adds $\mathcal{O}\left(\mathcal{B}\left(m^d\right)\right)$, resp. $\mathcal{O}\left(m^3\right)$.
Since these backsolves are the dominating factors in the cost of cg-iterations, the complexity of one cg-iteration step can be bounded by $\mathcal{O}\left(m^d \cdot \mathcal{B}\left((n/m)^d\right) + \mathcal{B}\left(m^d\right)\right)$ for general direct solvers, resp. by $\mathcal{O}\left(m^{-1} \cdot n^3 + m^3\right)$ for banded Cholesky solvers for 2D domains. These costs per iteration are the same for the conjugate gradient iterations with additive Schwarz preconditioners with linear and with multiscale interpolation.

(D) Overlook of the costs in the case of linear and of MsFE-coarsening: The upper bounds for the matrix sizes, setup costs of the preconditioners and costs of the cg-iterations are also shown in a short overview in Table 7.4. Using a banded Cholesky solver for 2D domains, this table simplifies to Table 7.5.

Task	**Add. Schwarz with linear interpolation**	**Add. Schwarz with MsFE interpolation**
Nodes of non-overl. subd.	$\mathcal{O}\left((n/m)^d\right)$	$\mathcal{O}\left((n/m)^d\right)$
Nodes of overlap. subd.	$\mathcal{O}\left((n/m)^d\right)$	$\mathcal{O}\left((n/m)^d\right)$
# coarse grid int. nodes	$\mathcal{O}\left(m^d\right)$	$\mathcal{O}\left(m^d\right)$
Setup local part of precond.	$\mathcal{O}\left(n^d\right)$	$\mathcal{O}\left(n^d\right)$
Setup coarse grid part of M^{-1}	$\mathcal{O}\left(n^d\right)$	$\mathcal{O}(n^d + m^d \cdot \mathcal{S}\left((n/m)^d\right))$
Cost of factorising local matrices	$\mathcal{O}\left(m^d \cdot \mathcal{S}\left((n/m)^d\right)\right)$	$\mathcal{O}\left(m^d \cdot \mathcal{S}\left((n/m)^d\right)\right)$
Cost of factorising global matrix	$\mathcal{O}\left(\mathcal{S}\left(m^d\right)\right)$	$\mathcal{O}\left(\mathcal{S}\left(m^d\right)\right)$
Overall setup cost	$\mathcal{O}(n^d + \mathcal{S}\left(m^d\right) + m^d \cdot \mathcal{S}\left((n/m)^d\right))$	$\mathcal{O}(n^d + \mathcal{S}\left(m^d\right) + m^d \cdot \mathcal{S}\left((n/m)^d\right))$
Cost of one cg-iteration	$\mathcal{O}(m^d \cdot \mathcal{B}\left((n/m)^d\right) + \mathcal{B}\left(m^d\right))$	$\mathcal{O}(m^d \cdot \mathcal{B}\left((n/m)^d\right) + \mathcal{B}\left(m^d\right))$

Table 7.4: Upper bounds for the computational complexity for additive Schwarz preconditioners with linear and with multiscale interpolation with minimal overlap using a general solver with cost $\mathcal{S}(r)$ for factorising and $\mathcal{B}(r)$ for the backsolve of a $r \times r$ matrix for the direct solves.

Task	**Add. Schwarz with linear interpolation**	**Add. Schwarz with MsFE interpolation**
Nodes of non-overl. subd.	$\mathcal{O}\left((n/m)^2\right)$	$\mathcal{O}\left((n/m)^2\right)$
Nodes of overlap. subd.	$\mathcal{O}\left((n/m)^2\right)$	$\mathcal{O}\left((n/m)^2\right)$
# coarse grid int. nodes	$\mathcal{O}\left(m^2\right)$	$\mathcal{O}\left(m^2\right)$
Setup local part of precond.	$\mathcal{O}\left(n^2\right)$	$\mathcal{O}\left(n^2\right)$
Setup coarse grid part of M^{-1}	$\mathcal{O}\left(n^2\right)$	$\mathcal{O}\left(m^{-2}n^4\right)$
Cost of factorising local matrices	$\mathcal{O}\left(m^{-2}n^4\right)$	$\mathcal{O}\left(m^{-2}n^4\right)$
Cost of factorising global matrix	$\mathcal{O}\left(m^4\right)$	$\mathcal{O}\left(m^4\right)$
Overall setup cost	$\mathcal{O}\left(m^{-2}n^4 + m^4\right)$	$\mathcal{O}\left(m^{-2}n^4 + m^4\right)$
Cost of one cg-iteration	$\mathcal{O}(m^{-1}n^3 + m^3)$	$\mathcal{O}(m^{-1}n^3 + m^3)$

Table 7.5: Upper bounds for the computational complexity for additive Schwarz preconditioners with linear and with multiscale interpolation with minimal overlap for $d = 2$ using a banded Cholesky solver for the direct solves.

From Table 7.5 we see that the overall setup cost can be minimised if $m = n^{2/3}$, the cost for the cg-iterations if $m = n^{3/4}$ (for minimal overlap, $d = 2$ and a banded Cholesky solver for the direct solves).

(E) Conclusions:
We see that the additional cost of setting up $A_{0,Ms}$ compared to A_0 is of the same order as factorising the local matrices. The MsFE-coarsening does therefore not increase the asymptotic complexity of the setup (but increases the multiplicative constant, which is hidden in the $\mathcal{O}$-notation). As the computation times in subsection 7.4 also show, for all numerical results considered, the extra cost of setting up the coarse grid matrices in the case of MsFE-interpolation can be neglected compared to the reduction in computation costs of the cg-iterations, when choosing this more sophisticated coarsening technique.

7.4 Computation times

Condition numbers and cg-iteration numbers are interesting in theory, but in practice one is mainly interested in a reduction of computation times by using good preconditioners. These (overall) computation times include both the time to set up the preconditioner (including the factorization of the local and coarse grid matrices) as well as the time needed to perform the cg-iterations. The times given in this section are for computations performed on a 1.5GHz Intel Pentium M processor with 496MB RAM. The code used was written in C++ (see Liberty [89]), using PetSc for the sparse matrices (see Balay et al. [8, 9]) and the Libmesh library (see Kirk et al. [83]) for the management of the meshes involved.

7.4.1 Two dimensional problems

Gridsize	Var.	$\lambda = 1/8$		$\lambda = 1/16$	
h	σ^2	Lin. Int.	Ms. Int.	Lin. Int.	Ms. Int.
1/64	1	$9(4+5)$	$10(5+5)$	$9(4+5)$	$10(5+5)$
	2	$9(4+5)$	$10(5+5)$	$10(4+6)$	$10(5+5)$
	4	$10(4+6)$	$10(5+5)$	$11(4+7)$	$11(5+6)$
	6	$10(4+6)$	$10(5+5)$	$12(4+8)$	$11(5+6)$
	8	$11(4+7)$	$11(5+6)$	$13(4+9)$	$12(5+7)$
	12	$12(4+8)$	$12(5+7)$	$16(4+12)$	$14(5+9)$
	16	$14(4+10)$	$12(5+7)$	$20(4+16)$	$15(5+10)$
	20	$15(4+11)$	$13(5+8)$	$24(4+20)$	$17(5+12)$
1/128	1	$21(5+16)$	$20(6+14)$	$21(5+16)$	$21(6+15)$
	2	$22(5+17)$	$21(6+15)$	$24(5+19)$	$23(6+17)$
	4	$25(5+20)$	$23(6+17)$	$31(5+26)$	$26(6+20)$
	6	$29(5+24)$	$25(6+19)$	$39(5+34)$	$29(6+23)$
	8	$33(5+28)$	$26(6+20)$	$49(5+44)$	$32(6+26)$
	12	$41(5+36)$	$30(6+24)$	$72(5+67)$	$40(6+34)$
	16	$52(5+47)$	$35(6+29)$	$103(5+98)$	$48(6+42)$
	20	$64(5+59)$	$39(6+33)$	$141(5+136)$	$60(6+54)$
1/256	1	$106(10+96)$	$98(12+86)$	$112(10+102)$	$105(12+93)$
	2	$122(10+112)$	$105(12+93)$	$138(10+128)$	$118(12+106)$
	4	$160(10+150)$	$121(12+109)$	$199(10+189)$	$146(12+134)$
	6	$212(10+202)$	$137(12+125)$	$279(10+269)$	$178(12+166)$
	8	$276(10+266)$	$166(12+154)$	$384(10+374)$	$210(12+198)$
	12	$436(10+426)$	$217(12+205)$	$679(10+669)$	$284(12+272)$
	16	$634(10+624)$	$271(12+259)$	$1,040(10+1,030)$	$364(12+352)$
	20	$912(10+902)$	$338(12+326)$	$1,453(10+1,443)$	$457(12+445)$

Table 7.6: Average Computation times in sec per field - Overall time (Setup time + Iteration time), $H = 1/16$ and overlap $\beta H = 1/64$

For $\Omega = (0,1)^2$ we obtain for $h = 1/64$, $h = 1/128$, resp. $h = 1/256$ and for $H = 1/16$ and overlap $\beta H = 1/64$ for different correlation lengths Tables 7.6 and 7.7. While for $h = 1/64$ the setup time is approximately of the same order as the time for the cg-iterations, it becomes rather neglectable for smaller h. Especially for $h = 1/128$ and $h = 1/256$ we find a significant reduction of computation times (up to a factor of 5) when using MsFE-interpolation instead of linear interpolation in the preconditioners.

Grids.	Var.	$\lambda = 1/32$		$\lambda = 1/64$	
h	σ^2	Lin. Int.	Ms. Int.	Lin. Int.	Ms. Int.
1/64	1	$9(4+5)$	$10(5+5)$	$9(4+5)$	$10(5+5)$
	2	$10(4+6)$	$10(5+5)$	$11(4+7)$	$11(5+6)$
	4	$12(4+8)$	$12(5+7)$	$14(4+10)$	$13(5+8)$
	6	$15(4+11)$	$13(5+8)$	$19(4+15)$	$15(5+10)$
	8	$18(4+14)$	$14(5+9)$	$25(4+21)$	$18(5+13)$
	12	$25(4+21)$	$16(5+11)$	$39(4+35)$	$22(5+17)$
	16	$34(4+30)$	$19(5+14)$	$61(4+57)$	$27(5+22)$
	20	$46(4+42)$	$24(5+19)$	$88(4+84)$	$36(5+31)$
1/128	1	$22(5+17)$	$22(6+16)$	$23(5+18)$	$22(6+16)$
	2	$27(5+22)$	$24(6+18)$	$28(5+23)$	$26(6+20)$
	4	$37(5+32)$	$29(6+23)$	$40(5+35)$	$32(6+26)$
	6	$52(5+47)$	$35(6+29)$	$57(5+52)$	$42(6+36)$
	8	$70(5+65)$	$40(6+34)$	$82(5+77)$	$50(6+44)$
	12	$121(5+116)$	$53(6+47)$	$153(5+148)$	$68(6+62)$
	16	$192(5+187)$	$69(6+63)$	$260(5+255)$	$86(6+80)$
	20	$286(5+281)$	$87(6+81)$	$425(5+420)$	$117(6+111)$
1/256	1	$112(10+102)$	$108(12+96)$	$116(10+106)$	$111(12+99)$
	2	$148(10+138)$	$121(12+109)$	$154(10+144)$	$130(12+118)$
	4	$215(10+205)$	$150(12+138)$	$234(10+224)$	$178(12+166)$
	6	$311(10+301)$	$185(12+173)$	$340(10+330)$	$242(12+230)$
	8	$461(10+451)$	$230(12+208)$	$532(10+522)$	$294(12+282)$
	12	$839(10+829)$	$303(12+291)$	$1,008(10+998)$	$428(12+416)$
	16	$1,639(10+1,638)$	$409(12+397)$	$2,320(10+2,310)$	$556(12+544)$
	20	$2,640(10+2,630)$	$562(12+550)$	$> 3,000$	$770(12+758)$

Table 7.7: Average Computation times in sec for 100 fields - Overall time (Setup time + Iteration time), $H = 1/16$ and overlap $\beta H = 1/64$

7.4.2 Three dimensional problems

We obtained similar results for small model problems on $\Omega = (0,1)^2$ with 24,576 fine grid elements, $H = 8h$ and $\beta H = 2h$, resp. $H = 4h$ and $\beta H = h$ for completely random fields. Again we find some reduction of computation times for fields with large variance σ^2. Here we chose linear boundary conditions for the construction of the coarse grid MsFE basis functions. Oscillatory boundary conditions and finer meshes are likely to reduce the number of cg-iterations even further (judging from the 2D results).

H	Overlap	Var. σ^2	Lin. int.	Ms/Lin.b.c.
8h	**2h**	**1**	$29(13+16)$	$30(14+16)$
		2	$29(13+16)$	$30(14+16)$
		4	$31(13+18)$	$30(14+16)$
		6	$31(13+18)$	$31(14+17)$
		8	$32(13+19)$	$31(14+17)$
		16	$34(13+21)$	$32(14+18)$
		32	$40(13+27)$	$36(14+22)$
4h	**h**	**1**	$15(7+8)$	$17(9+8)$
		2	$17(7+10)$	$17(9+8)$
		4	$21(7+14)$	$21(9+12)$
		6	$27(7+20)$	$25(9+16)$
		8	$33(7+26)$	$29(9+20)$
		16	$69(7+62)$	$53(9+44)$
		32	$181(7+174)$	$110(9+101)$

Table 7.8: Average computation times for 3D fields discretised with 24,576 fine grid elements and for 48 subdomains and overlap $\beta H = 2h$ (on the left), resp. 384 subdomains and overlap $\beta H = h$ (on the right) for completely random fields

7.5 Groundwater application (Mixed finite elements)

After these quite theoretical considerations we now would like to come back to the groundwater flow problem (see Chapter 2)

$$\alpha^{-1}\mathbf{v} + \boldsymbol{\nabla} u = 0, \tag{7.7}$$

$$\operatorname{div}(\mathbf{v}) = 0, \tag{7.8}$$

in two dimensions, with the simple domain $\Omega = (0,1)\times(0,1)$ and the following boundary conditions:

$$u = u_0 \neq 0 \text{ on } \{0\} \times (0,1), \tag{7.9}$$

$$u = 0 \text{ on } \{1\} \times (0,1), \tag{7.10}$$

and

$$\boldsymbol{\nabla} u.\boldsymbol{\nu} = 0 \text{ on } (0,1) \times \{0\} \text{ and } (0,1) \times \{1\}. \tag{7.11}$$

In Chapter 2 we discretised this problem to obtain a mixed finite element formulation, the saddle point problem

$$\begin{bmatrix} W(\alpha) & B \\ B^T & 0 \end{bmatrix} \begin{pmatrix} \mathbf{V} \\ \mathbf{U} \end{pmatrix} = \begin{pmatrix} \mathbf{G} \\ \mathbf{0} \end{pmatrix}. \tag{7.12}$$

Now let $\{\mathbf{z}_i\}$ denote a basis of $\ker(B^T)$ and let Z be the matrix with rows $\mathbf{z}_1^T, ..., \mathbf{z}_{\mathring{n}}^T$. Furthermore let $\mathring{A}(\alpha) := ZMZ^T$, let $\mathring{\mathbf{G}} := Z\mathbf{G}$ and let $\mathring{\mathbf{V}}$ be the vector such that $\mathbf{V} = Z^T\mathring{\mathbf{V}}$. Let $\mathcal{N}^h$ denote the nodes in the closed domain $\bar{\Omega}$, $\mathcal{N}_I^h$ the nodes in its interior, $\mathcal{N}_D^h$ the nodes on the Dirichlet boundary parts of $\bar{\Omega}$ and let $\mathcal{N}_{N,1}^h, ..., \mathcal{N}_{N,s_N}^h$ denote the sets of nodes in the maximum sized connected Neumann boundary parts $\Gamma_N^1, ..., \Gamma_N^{s_N}$ of $\bar{\Omega}$, with $\Gamma_N = \Gamma_N^1 \cup ... \cup \Gamma_N^{s_N}$, $\Gamma_N^i \cap \Gamma_N^j = \emptyset$ for all $1 \leq i \neq j \leq s_N$. Note that if a node is an end point of both a Dirichlet and a Neumann boundary, we consider it to be a Neumann boundary node. Let ϕ_i for i with $\mathbf{x}_i^h \in \mathcal{N}^h\left(\bar{\Omega}\right)$ be the basis of head functions with $\phi_i(\mathbf{x}_j^h) = \delta_{ij}$ as defined in the beginning of Chapter 4 and illustrated in Figure 4-1. With this basis let

$$\mathcal{A}_{ij}(\alpha) := \int_\Omega \alpha^{-1}(\mathbf{x})\boldsymbol{\nabla}\phi_i.\boldsymbol{\nabla}\phi_j d\mathbf{x}, \qquad \text{for all } \mathbf{x}_i^h, \mathbf{x}_j^h \in \mathcal{N}^h. \tag{7.13}$$

Then define a matrix A with entries

$$A_{ij}(\alpha) := \mathcal{A}_{ij}(\alpha), \qquad \text{for all } \mathbf{x}_i^h, \mathbf{x}_j^h \in \mathcal{N}_I^h \cup \mathcal{N}_D^h. \tag{7.14}$$

Moreover define matrices C and D with

$$C_{ik} := \sum_{\mathbf{x}_j^h \in \mathcal{N}_{N,k}^h} \mathcal{A}_{ij}, \qquad \text{for all } \mathbf{x}_i^h \in \mathcal{N}_I^h \cup \mathcal{N}_D^h \quad \text{and} \quad k = 1, ..., s_N, \tag{7.15}$$

and

$$D_{kl} := \sum_{\mathbf{x}_i^h \in \mathcal{N}_{N,k}^h} \sum_{\mathbf{x}_j^h \in \mathcal{N}_{N,l}^h} \mathcal{A}_{ij}, \qquad \text{for all } k, l = 1, ..., s_N. \tag{7.16}$$

Based on these definitions we can then decouple the velocity vector $\mathbf{V}$ from the pressure vector $\mathbf{U}$ in equation (7.12) to obtain the following proposition as shown in Scheichl [116], Proposition 4.2.

PROPOSITION 7.1

Let Ω be simply connected and $\Gamma_N \neq \emptyset$. Then

$$\mathring{A}(\alpha)\mathring{\mathbf{V}} = \begin{bmatrix} A(\alpha) & C(\alpha) \\ C^T(\alpha) & D(\alpha) \end{bmatrix} \mathring{\mathbf{V}} = \mathring{\mathbf{G}}. \tag{7.17}$$

Let us now come back to the simple model problem on a square domain with Neumann boundary conditions on two opposite sides of the domain and Dirichlet conditions on the other two sides (as given in Figure 2-1). Since Γ_N in our model problem therefore only contains two components Γ_N^1 and Γ_N^2, system (7.17) simplifies to

$$\begin{bmatrix} A(\alpha) & \mathbf{c}(\alpha) \\ \mathbf{c}^T(\alpha) & d(\alpha) \end{bmatrix} \mathring{\mathbf{V}} = \mathring{\mathbf{G}}, \tag{7.18}$$

where $A(\alpha)$ is a square sparse matrix, $\mathbf{c}(\alpha)$ a single column vector and $d(\alpha)$ a scalar. Rewriting the vectors $\mathring{\mathbf{V}} = \left[\mathring{\boldsymbol{\eta}}; \mathring{\rho}\right]$ and $\mathring{\mathbf{G}} = \left[\mathring{\boldsymbol{\varphi}}; \mathring{\zeta}\right]$ with scalars $\mathring{\rho}, \mathring{\zeta}$ and vectors $\mathring{\boldsymbol{\eta}}, \mathring{\boldsymbol{\varphi}}$ and one can apply a block decomposition technique reducing (7.18) to the following two linear systems, which we have to solve for vectors $\mathbf{x}$ and $\mathbf{y}$

$$A\mathbf{x} = \mathbf{c}, \qquad A\mathbf{y} = \mathring{\boldsymbol{\varphi}}. \tag{7.19}$$

Then $\mathring{\rho}$ is simply given by

$$\mathring{\rho} = \left(\mathring{\zeta} - \mathbf{c}^T\mathbf{y}\right) / \left(d - \mathbf{c}^T\mathbf{x}\right), \tag{7.20}$$

from which one obtains

$$\mathring{\boldsymbol{\eta}} = \mathbf{y} - \mathring{\rho}\mathbf{x}. \tag{7.21}$$

It turns in fact out that for our model problem in which the Dirichlet data is constant on each component of Γ_D, we obtain $\mathring{\boldsymbol{\varphi}} = 0$.
Therefore we get $\mathbf{y} = 0$ as well as

$$\mathring{\rho} = \mathring{\zeta} / \left(d - \mathbf{c}^T\mathbf{x}\right)$$

and

$$\mathring{\boldsymbol{\eta}} = -\mathring{\rho}\mathbf{x} = -\left[\mathring{\zeta} / \left(d - \mathbf{c}^T\mathbf{x}\right)\right]\mathbf{x}.$$

All that remains to be done to solve (7.17) is to solve the linear system

$$A\mathbf{x} = \mathbf{c}. \tag{7.22}$$

As shown in Scheichl [116], Subsection 4.2.1, this system is exactly of the form of the systems considered earlier in this thesis and we can therefore test the newly developed additive Schwarz preconditioners with multiscale coarsening on this groundwater flow problem.

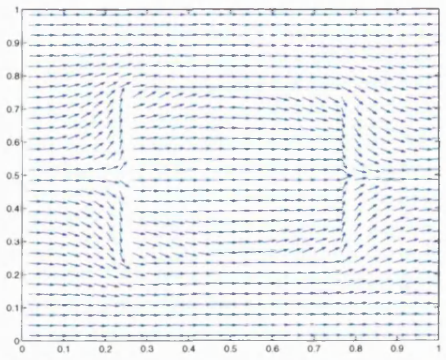
Let us now define the following:

DEFINITION 7.2
For a ***velocity field*** *of a moving fluid we associate a velocity vector to each point in the fluid.*
In fluid dynamics, ***streamlines*** *are a family of curves that are instantaneously tangent to the velocity vector of the flow.*

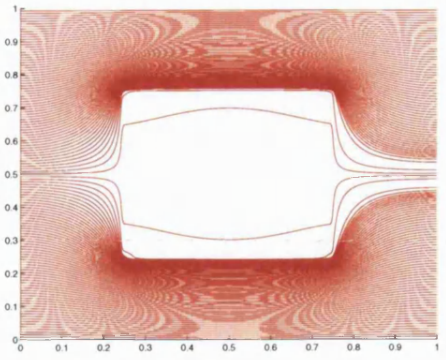
We now finally consider plots of the velocity fields and streamlines for several permeability fields.

The first two experiments are for two-media problems with $\alpha = 1$ and $\alpha = \hat{\alpha}$, $\hat{\alpha} \neq 1$, in different parts of Ω. In Figure 7-2 we choose $\alpha = \hat{\alpha}$ inside a square with corners $(0.25, 0.25)$, $(0.75, 0.25)$, $(0.25, 0.75)$ and $(0.75, 0.75)$ and $\alpha = 1$ everywhere else in the unit square domain $(0, 1) \times (0, 1)$. In Figure 7-3 we choose a field with $\alpha = \hat{\alpha}$ inside the triangle with corners $(1.0, 0.0)$, $(1.0, 1.0)$ and $(0.5, 0.5)$ and with $\alpha = 1$ elsewhere.

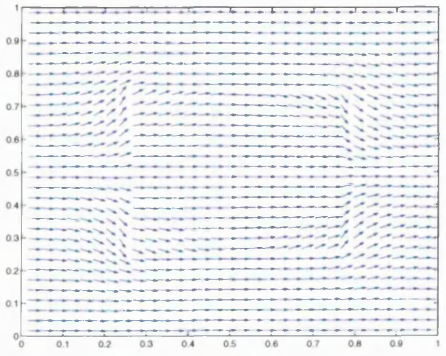
Last but not least we give the plots for realisations of random fields of Ornstein-Uhlenbeck type with correlation length $\lambda = 4h$, resp. $\lambda = 10h$ for several values of σ^2 (see plots of the permeability fields in Figure 7-4 and of the velocity fields and streamlines in Figure 7-5, resp. Figure 7-6). With increased variance σ^2 we observe (locally) stronger varying vector fields.



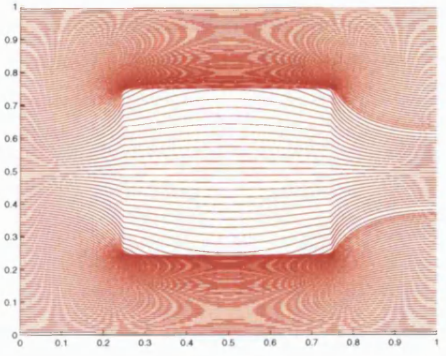
(a) Velocity field - $\hat{\alpha} = 0.001$



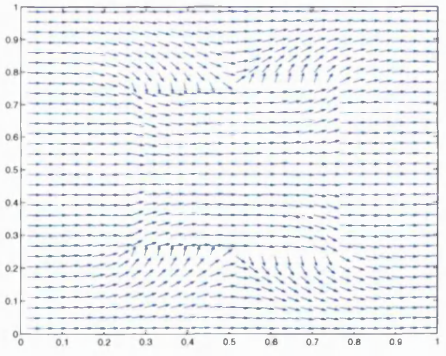
(b) Streamlines - $\hat{\alpha} = 0.001$



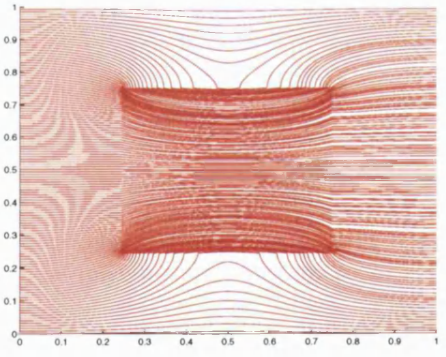
(c) Velocity field - $\hat{\alpha} = 0.1$



(d) Streamlines - $\hat{\alpha} = 0.1$

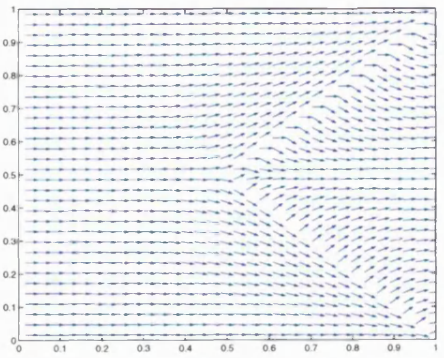


(e) Velocity field - $\hat{\alpha} = 1000$

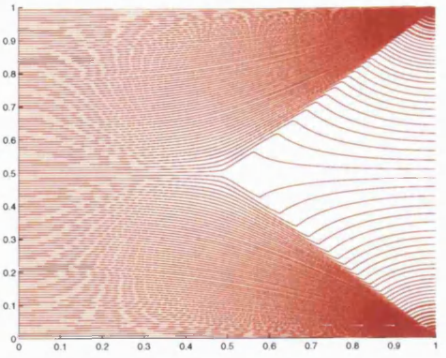


(f) Streamlines - $\hat{\alpha} = 1000$

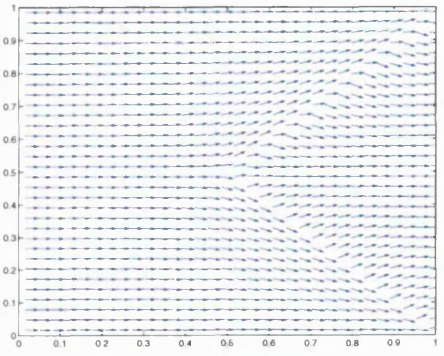
Figure 7-2: Vector plots and streamlines for a field with $\alpha = \hat{\alpha}$ in the square $[0.25, 0.75] \times [0.25, 0.75]$ at the centre of the domain and with $\alpha = 1$ elsewhere.



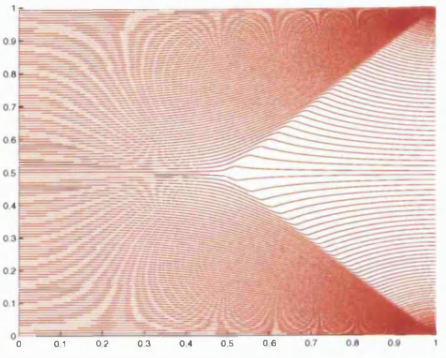
(a) Velocity field - $\hat{\alpha} = 0.001$



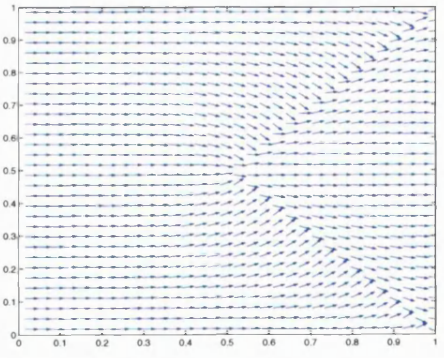
(b) Streamlines - $\hat{\alpha} = 0.001$



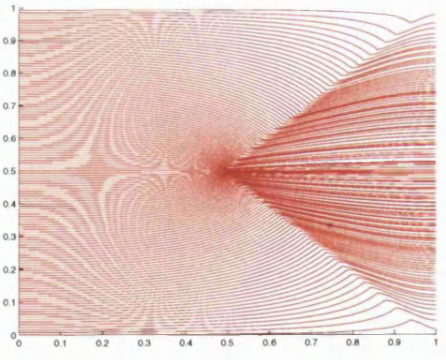
(c) Velocity field - $\hat{\alpha} = 0.1$



(d) Streamlines - $\hat{\alpha} = 0.1$

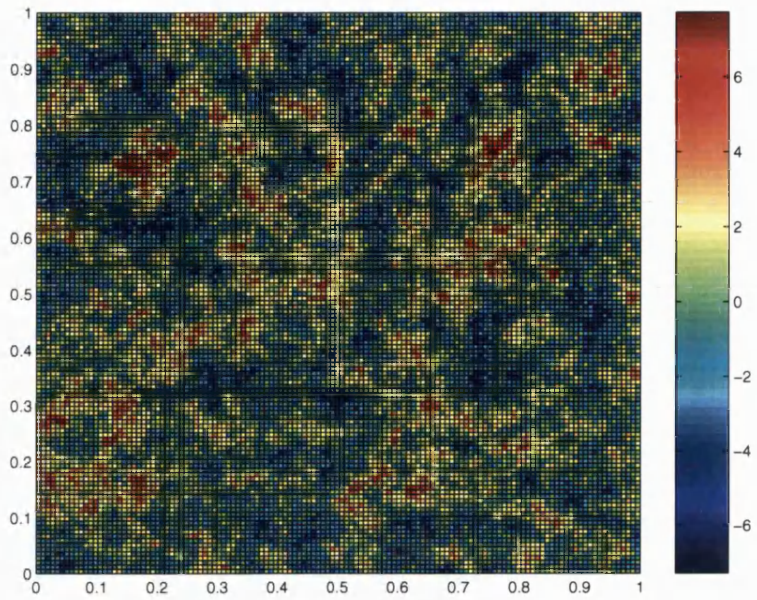


(e) Velocity field - $\hat{\alpha} = 1000$

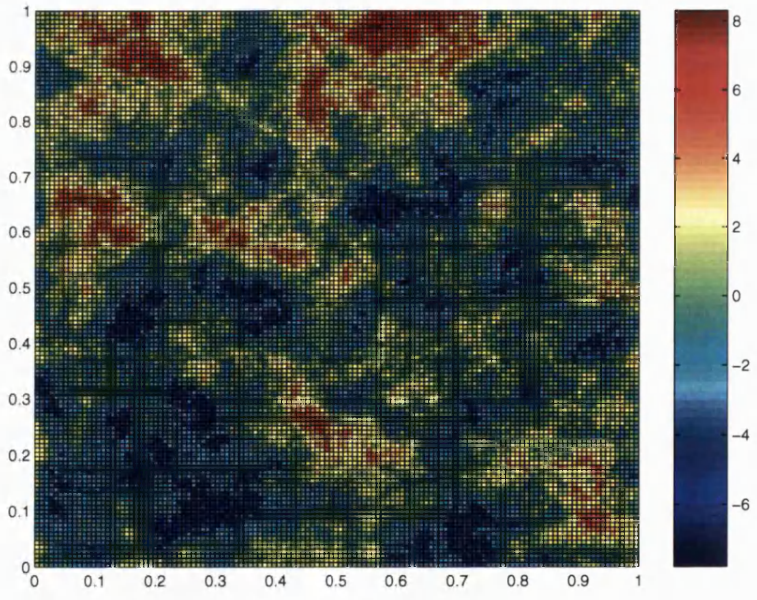


(f) Streamlines - $\hat{\alpha} = 1000$

Figure 7-3: Vector plots and streamlines for a field with $\alpha = \hat{\alpha}$ in the triangle with corners $(1.0, 0.0)$, $(1.0, 1.0)$ and $(0.5, 0.5)$ and with $\alpha = 1$ elsewhere.

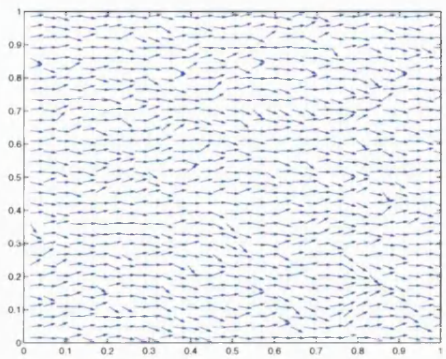


(a) $\lambda = 4h$

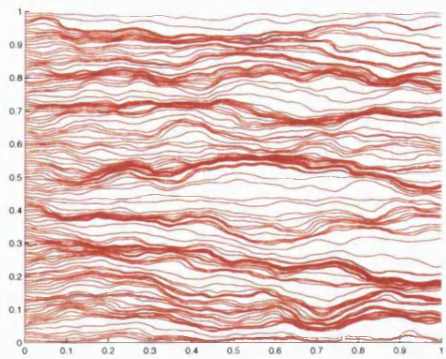


(b) $\lambda = 10h$

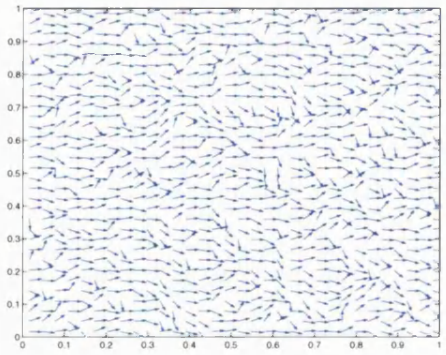
Figure 7-4: Realisations of Gaussian random fields with covariance of Ornstein-Uhlenbeck type with $h = 1/128$, $\sigma^2 = 4$ and $\lambda = 4h$, resp. $\lambda = 10h$.



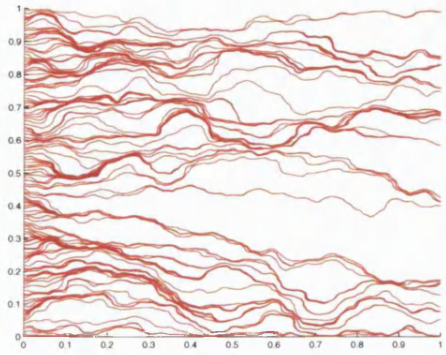
(a) Velocity field - $\sigma^2 = 1$



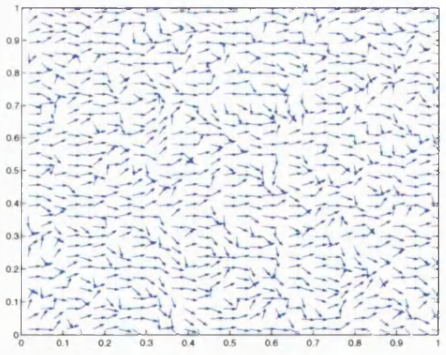
(b) Streamlines - $\sigma^2 = 1$



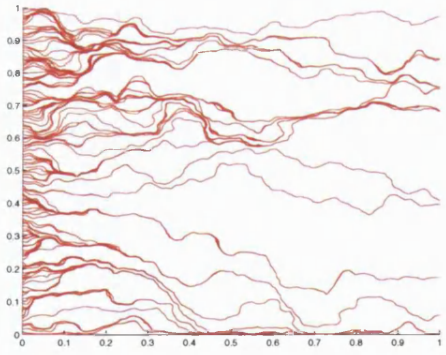
(c) Velocity field - $\sigma^2 = 4$



(d) Streamlines - $\sigma^2 = 4$

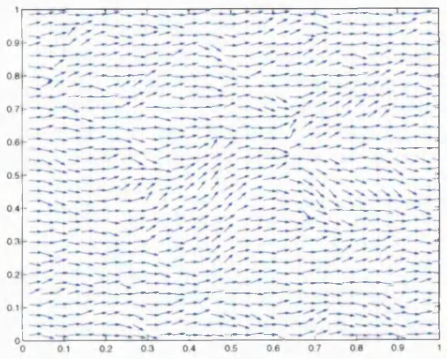


(e) Velocity field - $\sigma^2 = 8$

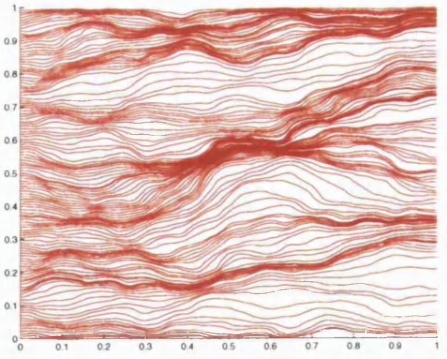


(f) Streamlines - $\sigma^2 = 8$

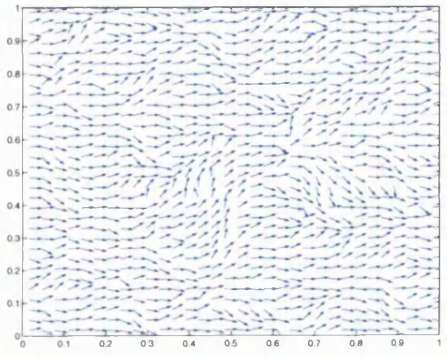
Figure 7-5: Vector plots and streamlines for the realisation of a random field given above with $h = 1/128$, $\lambda = 4h$ and $\sigma^2 = 1$, 4, resp. 8.



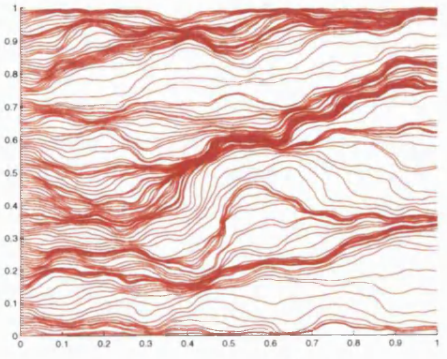
(a) Velocity field - $\sigma^2 = 1$



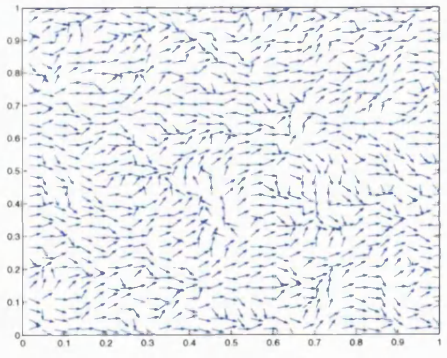
(b) Streamlines - $\sigma^2 = 1$



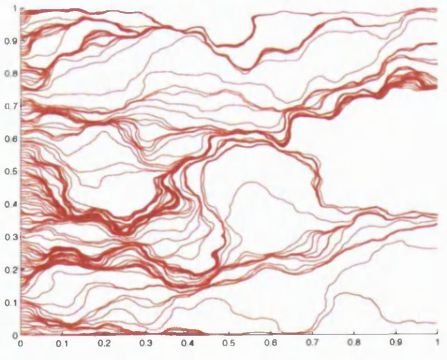
(c) Velocity field - $\sigma^2 = 4$



(d) Streamlines - $\sigma^2 = 4$



(e) Velocity field - $\sigma^2 = 8$



(f) Streamlines - $\sigma^2 = 8$

Figure 7-6: Vector plots and streamlines for the realisation of a random field given above with $h = 1/128$, $\lambda = 10h$ and $\sigma^2 = 1, 4,$ resp. 8.

7.6 Summary

In this last chapter we first of all gave a short overview of the theoretical bounds, that were developed in the previous chapters. The theoretical results were then supported by a series of numerical results on the cg-iteration numbers for 2D and 3D problems with underlying random permeability fields. These showed how significantly multiscale interpolation can improve the convergence rate compared to linear coarsening and that the skeleton based preconditioners are in fact exact inverses of $A(\alpha)$ (in the case of zero overlap). Next we developed a complexity analysis which studied both the setup cost of the different two-level additive Schwarz preconditioners and the complexity of one cg-iteration step when using these preconditioners.
More important than condition numbers and cg-iteration numbers in practice however are computation times. We therefore studied these times for various random fields in 2D and 3D to show that multiscale coarsening can really help to increase the speed of solvers and can therefore in many cases significantly save costs in practice.
Finally we finished the presentation of new results by applying our preconditioners to the groundwater flow problem (7.7, 7.8). We gave streamline and vector field plots for both simple two-media and random field problems and discussed these results.

7.7 Concluding remarks

Finally we would like to give a short overview of what we think might be the most important contributions of this thesis and how they might result in interesting new projects in the future.

In Chapters 2 to 4 of the thesis we proved that groundwater flow problems can be very ill-conditioned due to underlying permeability fields. We then showed how random fields of Ornstein-Uhlenbeck type can be used to model the permeability and how the condition number of the resulting stiffness matrix when discretising with the help of finite elements depends on the parameters of these fields, their variance, mean and correlation length. By combining results from stochastics, linear algebra and numerical analysis we showed, how ideas from different fields of mathematics can sometimes help other areas to get new important insights. Poisson clumping heuristics to our best knowledge for example were never used for the analysis of multiscale problems before. We gave sharp bounds for the maxima of random fields and of global and local ratios of these fields and showed how these bounds can be used in practice to help to estimate expected computation times when using a large number of fields. This might aid to planning security for researchers, companies, etc. that model with the help of random fields. Although there is no closed formula for probability distributions of general random fields, similar bounds to those derived in Chapter 3 can be found for fields with different covariance and we hope that our analysis might inspire others to do more research in this area.

In Chapter 5 we then focused on a first class of preconditioners for the groundwater flow problem, the two-level additive Schwarz preconditioners with linear interpolation. While previous research mainly discussed the case in which the coarse grid resolves the permeability field, we now tried to work out how the condition number of the preconditioned matrix depends on the underlying field, if we cannot or do not want to adapt the coarse grid to the underlying field. We therefore gave sharp upper bounds for the condition number for deterministic permeability as well as for random permeability fields of Ornstein-Uhlenbeck type.

Since these theoretical and numerical results show that preconditioning with additive Schwarz preconditioners with linear interpolation can still lead to very ill-conditioned systems especially if the underlying field has got short correlation length and large variance, we then tried to find better coarsening operators. These are based on the

multiscale finite element method and we showed that this method cannot only be used to increase the accuracy of computations on coarse meshes, but that it can also be applied to speed up computation times. We used multiscale finite elements for the first time as part of an additive Schwarz preconditioner and discussed the importance of the boundary conditions for the construction of the multiscale basis functions. The method was analysed for some simple model problems. A more detailed discussion will be given in Graham et al. [60]. We showed that in certain cases MsFE-coarsening can achieve independence of the convergence of the preconditioned system on the permeability field, if the field is constant in the overlap of the subdomains. Furthermore we gave some first idea, why the size of the overlap can be crucial for the convergence. A similar analysis could be done in future projects for the multiplicative and hybrid Schwarz preconditioners, that we have introduced in Subsection 5.1.4. We analysed the complexity of the setup of the preconditioners described in this thesis as well as the cost of applying them within a conjugate gradient solver. Next we discussed why these algorithms can be very well parallelised. With respect to this classification it might be interesting to find out more about the possible preconditioners that one could find with complexities between the node- and skeleton-based additive Schwarz preconditioners with multiscale coarsening. Also the possibilities that MsFE-coarsening could have for non-overlapping domain decomposition methods like Bramble-Pasciak-Schwarz preconditioners (see Bramble et al. [15]) or vertex space preconditioners (see Smith [121, 122]) should be mentioned in this context.

The numerical results for 1D, 2D and 3D domains for two media problems as well as for problems with underlying permeability fields, that we gave in Chapters 6 and 7, supported the theory and gave an idea of the potential of MsFE-coarsening for groundwater flow problems. While this class of problems allows a detailed analysis, it seems that interpolation with the help of multiscale basis functions could also be applied to a much larger range of problems. In fact for many problems with strongly varying coefficients the newly developed techniques might mean significant improvements in computation times. Multiscale coarsening methods therefore seem to be a tool, that could provide ideas for many areas of research and we hope that this thesis can help to see their potential.

Appendix A

Some basic results on eigenvalues and eigenvectors

In this appendix we would like to give some definitions and simple lemmata on eigenvalues and eigenvectors that are used throughout this thesis (for proofs see for example Golub and Van Loan [58]).

DEFINITION A.1

(i) For a given matrix $A \in \mathbb{R}^{n\times n}$ *its* ***eigenvalues*** $\lambda \in \mathbb{C}$ *and* ***eigenvectors*** $\mathbf{v} \in \mathbb{C}^n \setminus \{0\}$ *are the solution of*

$$A\mathbf{v} = \lambda\mathbf{v}.$$

(ii) The set of eigenvalues of A *are called the matrix'* ***spectrum*** *and is denoted by* $\sigma(A)$.

(iii) The ***spectral radius*** *is defined as*

$$\rho(A) := \max_{\lambda\in\sigma(A)} \{|\lambda|\}.$$

(iv) If $\|\cdot\|$ *denotes a matrix norm, the* ***condition number*** *of an invertible matrix* A *is defined to be*

$$\kappa(A) := \|A\| \left\|A^{-1}\right\|.$$

(v) Given a second matrix $M \in \mathbb{R}^{n\times n}$ *the* ***generalized eigenproblem*** *of the two matrices is finding eigenvalues* λ *and eigenvectors* $\mathbf{v}$ *such that*

$$A\mathbf{v} = \lambda M\mathbf{v}.$$

(vi) *If all its eigenvalues have positive real part, then A is called* **positive definite**.

(vii) *For an arbitrary matrix B, let*

$$\|B\|_A := \sup_{\mathbf{v}\in\mathbb{R}^n} \frac{\|B\mathbf{v}\|_A}{\|\mathbf{v}\|_A}$$

with $\|\mathbf{v}\|_A^2 := \mathbf{v}^T A\mathbf{v}$.

With these definitions the following statements follow.

LEMMA A.2
Let $A, M \in \mathbb{R}^{n\times n}$ be two symmetric positive definite matrices. Then we obtain:

(i) *The following eigenvalue problems have the same eigenvalues*

$$A\mathbf{v} = \lambda M\mathbf{v}, \tag{A.1}$$
$$M^{-1}A\mathbf{v} = \lambda\mathbf{v}, \tag{A.2}$$
$$\left(M^{-1/2}AM^{-1/2}\right)\mathbf{v} = \lambda\mathbf{v}, \tag{A.3}$$
$$\left(A^{1/2}M^{-1}A^{1/2}\right)\mathbf{v} = \lambda\mathbf{v}. \tag{A.4}$$

The eigenvalues are all real and positive.

(ii) *The minimum eigenvalue $\lambda_{\min}$ and maximum eigenvalue $\lambda_{\max}$ of the 4 problems above satisfy*

$$\lambda_{\min} = \inf_{\mathbf{v}\in\mathbb{R}^n} \frac{\mathbf{v}^T A\mathbf{v}}{\mathbf{v}^T M\mathbf{v}},\ \lambda_{\max} = \sup_{\mathbf{v}\in\mathbb{R}^n} \frac{\mathbf{v}^T A\mathbf{v}}{\mathbf{v}^T M\mathbf{v}}. \tag{A.5}$$

(iii) *It is*

$$\left\|M^{-1}A\right\|_A = \left\|M^{-1}A\right\|_M = \lambda_{\max} = \rho\left(M^{-1}A\right), \tag{A.6}$$
$$\left\|\left(M^{-1}A\right)^{-1}\right\|_A = \left\|\left(M^{-1}A\right)^{-1}\right\|_M = 1/\lambda_{\min}, \tag{A.7}$$

and therefore

$$\kappa_A\left(M^{-1}A\right) = \kappa_2\left(M^{-1/2}AM^{-1/2}\right) = \lambda_{\max}/\lambda_{\min}. \tag{A.8}$$

Appendix B

The Preconditioned Conjugate Gradient Method

The Preconditioned Conjugate Gradient Algorithm (PCG) can be used to solve symmetric, positive definite (spd) equation systems. Consider therefore the spd system

$$A\mathbf{x} = \mathbf{b},\ A \in \mathbb{R}^{n\times n},\ \mathbf{x}, \mathbf{b} \in \mathbb{R}^n. \tag{B.1}$$

In the groundwater flow problem matrix A is typically ill-conditioned due to the behaviour of the underlying PDE. Let us now assume that we can find a spd preconditioner M, which approximates A in the sense that the condition number of $M^{-1}A$ is significantly smaller than that of A. Furthermore, for practical reasons, we require for M, that $M\mathbf{x} = \mathbf{b}$ is inexpensive to solve, as in the PCG algorithm a linear system with matrix M has to be solved in each step. The preconditioned system is then given by

$$M^{-1}A\mathbf{x} = M^{-1}\mathbf{b},\ M, A \in \mathbb{R}^{n\times n},\ \mathbf{x}, \mathbf{b} \in \mathbb{R}^n. \tag{B.2}$$

$M^{-1}A$ is not necessarily spd, if A and M are spd, but for the algorithm to work, we will have to make sure that this is still the case.
We can then formulate the following version of the **Preconditioned Conjugate Gradient Algorithm**, which is suitable for implementation on a (serial or parallel) computer:

(i) *Choose an initial guess* $\mathbf{x}_0$.

(ii) *Compute* $\mathbf{r}_0 := \mathbf{b} - A\mathbf{x}_0$, $\mathbf{z}_0 := M^{-1}\mathbf{r}_0$ *and* $\mathbf{p}_0 := \mathbf{z}_0$.

(iii) ***For*** $k = 0, 1, \ldots,$ ***until*** *convergence* ***do:***

(iv) $\alpha_k := (\mathbf{r}_k, \mathbf{z}_k) \,/\, (A\mathbf{p}_k, \mathbf{p}_k)$

(v) $\mathbf{x}_{k+1} := \mathbf{x}_k + \alpha_k \mathbf{p}_k$

(vi) $\mathbf{r}_{k+1} := \mathbf{r}_k - \alpha_k A\mathbf{p}_k$

(vii) $\mathbf{z}_{k+1} := M^{-1}\mathbf{r}_{k+1}$

(viii) $\beta_j := (\mathbf{r}_{k+1}, \mathbf{z}_{k+1}) \,/\, (\mathbf{r}_k, \mathbf{z}_k)$

(ix) $\mathbf{p}_{k+1} := \mathbf{z}_{k+1} + \beta_k \mathbf{p}_k$

(x) ***end do.***

Important features of the method:

- The *cost* of each iteration step is dominated by the cost of the matrix vector multiplication with A (in (iv) and (vi)) and the solution of the linear equation with M in (vii). Furthermore we only need to compute four dot products, three scale operations, three vector-vector additions/subtractions and two scalar divisions.

- The *storage requirements* for the algorithm are only four vectors of length n ($\mathbf{z}$,$\mathbf{p}$,$\mathbf{u}$ and $\mathbf{r}$), two scalars (α and β) and whatever storage is needed to solve $\mathbf{z} = M^{-1}\mathbf{r}$.

- The algorithm can be *parallelised* effectively. This almost optimal parallelisation includes the preconditioner, if we choose an additive Schwarz preconditioner, as shown in Chapter 6.

- The matrices A and M *appear only through their action.* Therefore we do not have to set up A and M explicitly, e.g. we can have A in form of element or local stiffness matrices.

- The *convergence rate* of the PCG method is directly related to the condition

number of the preconditioned system by

$$\|\mathbf{x}-\mathbf{x}_k\|_A \leq 2\left(\frac{\sqrt{\kappa\left(M^{-1}A\right)}-1}{\sqrt{\kappa\left(M^{-1}A\right)}+1}\right)^k \|\mathbf{x}-\mathbf{x}_0\|_A \tag{B.3}$$

where $\|\cdot\|_A$ is the norm induced by A (see Appendix A). This is also the main motivation of using preconditioners in the first place, as this means that if M approximates A well enough and can be inverted without large extra cost, the overall amount of work can be reduced.

- *Stopping criterion:* Let us suppose that the PCG method runs until the relative error satisfies a pre-selected tolerance, θ, i.e. $\|\mathbf{x}-\mathbf{x}_k\|_A / \|\mathbf{x}-\mathbf{x}_0\|_A \leq \theta$. With the previous observation, this tolerance satisfies the inequality

$$\theta \geq 2\left(\frac{\sqrt{\kappa\left(M^{-1}A\right)}-1}{\sqrt{\kappa\left(M^{-1}A\right)}+1}\right)^k = 2\left(1-\frac{2/\sqrt{\kappa\left(M^{-1}A\right)}}{1+1/\sqrt{\kappa\left(M^{-1}A\right)}}\right)^k.$$

If we now take the logarithm on both sides of this inequality and use the fact that $\log\left(1-\frac{2/\sqrt{\kappa(M^{-1}A)}}{1+1/\sqrt{\kappa(M^{-1}A)}}\right) \approx -2/\sqrt{\kappa\left(M^{-1}A\right)}$ for large $\kappa\left(M^{-1}A\right)$, we obtain an upper bound for the number of PCG-iterations

$$k \lesssim \frac{1}{2}\left|\log\left(\theta/2\right)\right| \cdot \sqrt{\kappa\left(M^{-1}A\right)},$$

i.e. it grows at most with the square root of the condition number.

Appendix C

List of Notations

Symbol		Definition	Page
Chapter 1:			
μ	:=	dynamic viscosity	1
α	:=	permeability	1
$\mathbf{v}$	:=	velocity	1
u	:=	pressure	1
Ω	:=	open, bounded and connected subset of $\mathbb{R}^d$	1
$\mathbf{x}$	:=	vector in $\Omega \subset \mathbb{R}^d$	1
$\Gamma = \Gamma^D \cup \Gamma^N$	:=	boundary of Ω, partitioned into its Dirichlet part, Γ^D, and its Neumann part, Γ^N.	1
$\boldsymbol{\nu}$	:=	outward unit normal from Ω at $\mathbf{x} \in \Gamma$	1
$\mathbf{U}$, $\mathbf{V}$	:=	approximations of u and $\mathbf{v}$ in finite dimensional spaces	2
$W(\alpha)$	:=	mass matrix arising from the discretisation of the operator $\mathbf{v} \longmapsto \mu\alpha^{-1}\mathbf{v}$	2
B^T	:=	discrete divergence operator	2
n_h	:=	number of degrees of freedom in Ω	2
$A(\alpha)$	:=	symmetric positive definite $n_h \times n_h$ stiffness matrix	2
$\Sigma(\mathbf{x}, \mathbf{y})$	:=	covariance for random field at points $\mathbf{x}, \mathbf{y} \in \Omega$	3
λ	:=	correlation length of a random field	3
$\kappa(A(\alpha))$	:=	condition number of $A(\alpha)$	3
M^{-1}	:=	preconditioner of $A(\alpha)$	4
Ω_j	:=	open, non-overlapping subdomains of Ω, $j = 1, ..., p$	4
$\widetilde{\Omega_j}$	:=	open, overlapping subdomains of Ω, $j = 1, ..., p$	4

Symbol		Definition	Page
Chapter 2:			
v	:=	specific discharge / Darcy velocity	9
u_R	:=	residual pressure	9
z	:=	fluid height	9
ρ	:=	density	9
g	:=	acceleration	9
$L_2(\Omega)$	:=	$\{u: \quad \Omega \to \mathbb{R} \quad \mid \quad \int_\Omega u^2 < \infty\}$	13
$H^1(\Omega)$	:=	$\left\{u: \quad \Omega \to \mathbb{R} \quad \mid \quad u, \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y} \in L_2(\Omega)\right\}$	13
$\mathcal{V}$	:=	$\{v: \quad \Omega \to \mathbb{R} \quad \mid \quad v \in H^1(\Omega), \quad v(\mathbf{x}) = 0 \text{ for } \mathbf{x} \in \Gamma_D\}$	13
$\mathcal{V}_g$	:=	$\{v: \quad \Omega \to \mathbb{R} \quad \mid \quad v \in H^1(\Omega), \quad v(\mathbf{x}) = g \text{ for } \mathbf{x} \in \Gamma_D\}$	13
$(u,v)_{H^1(\Omega),\alpha}$	:=	$\int_\Omega \alpha(\mathbf{x}) \nabla u(\mathbf{x}) . \nabla v(\mathbf{x}) d\mathbf{x}$	13
$L(v)$	:=	$\int_\Omega f(\mathbf{x}) v(\mathbf{x}) d\mathbf{x} + \int_{\Gamma_N} g(\mathbf{x}) v(\mathbf{x}) d\mathbf{x}$	13
$\mathcal{T}^h(\bar{\Omega})$	:=	triangulation of Ω with closed elements τ_i^h of diameter h	14
$\overset{\circ}{\tau}_i^h$	:=	interior of (fine) grid element τ_i^h	14
$\mathcal{V}^h$	:=	$\{v_h \in C(\Omega) \quad \mid \quad v_h\|_\tau \text{ linear } \forall \tau \in \mathcal{T}^h(\bar{\Omega}) \text{ and } v = 0 \text{ on } \Gamma_D\}$	15
$\mathcal{V}_g^h$	:=	$\{v_h \in C(\Omega) \quad \mid \quad v_h\|_\tau \text{ linear } \forall \tau \in \mathcal{T}^h(\bar{\Omega}) \text{ and } v = g \text{ on } \Gamma_D\}$	15
$\mathcal{N}^h(\bar{\Omega})$	:=	set of nodes $\mathbf{x}_i^h$ of the triangulation $\mathcal{T}^h(\Omega)$	15
$\delta_{i,j}$	:=	Kronecker delta	16
A_{ij}	:=	$(\phi_i, \phi_j)_{H^1(\Omega),\alpha}$ with (fine grid) hat functions ϕ_i, ϕ_j	16
f_i	:=	$L(\phi_i)$ with fine grid hat function ϕ_i	16
$H(\text{div}, \Omega)$	:=	$\left\{\mathbf{w} \in (L_2(\Omega))^d: \quad \text{div}(\mathbf{w}) \in L_2(\Omega)\right\}$	17
$(\mathbf{v}, \mathbf{w})_{H(\text{div},\Omega)}$	:=	$\int_\Omega (\mathbf{v}.\mathbf{w} + \text{div}(\mathbf{v}) \cdot \text{div}(\mathbf{w}))\, d\mathbf{x}$	17
$\|u\|_{L_2(\Omega)}$	:=	$\int_\Omega u^2(\mathbf{x}) d\mathbf{x}$	17
$\mathcal{R}$	:=	$H_{0,N}(\text{div}, \Omega) \qquad (:= \{\mathbf{v} \in H(\text{div}, \Omega) \quad \mid \quad \mathbf{v}.\boldsymbol{\nu} = 0 \text{ on } \Gamma_N\})$	17
$\mathcal{S}$	:=	$L_2(\Omega)$	17
$\langle g, \mathbf{r}.\boldsymbol{\nu} \rangle_{L_2(\Gamma_D)}$	:=	$\int_{\Gamma_D} g \mathbf{r}.\boldsymbol{\nu} d\mathbf{x}$	18
$\mathcal{R}_h$, $\mathcal{S}_h$	:=	finite dimensional subspaces of $\mathcal{R}$ and $\mathcal{S}$	18
$w(\mathbf{v}, \mathbf{r})$	:=	$\left(\alpha^{-1}\mathbf{v}, \mathbf{r}\right)_{L_2(\Omega)}$	18
$b(\mathbf{r}, s)$	:=	$-(\text{div}(\mathbf{v}), s)_{L_2(\Omega)}$	18
$G(\mathbf{r})$	:=	$\langle g, \mathbf{r}.\boldsymbol{\nu} \rangle_{L_2(\Gamma_D)}$	18
n_r, n_s	:=	dimension of $\mathcal{R}_h$, resp. of $\mathcal{S}_h$	19
$\{\mathbf{r}_1, ..., \mathbf{r}_{n_r}\}$	:=	basis of $\mathcal{R}_h$	19
$\{s_1, ..., s_{n_s}\}$	:=	basis of $\mathcal{S}_h$	19

Symbol		Definition	Page
$W_{i,j}(\alpha)$	:=	$w(\mathbf{r}_i, \mathbf{r}_j)$ (:= entry of Raviart-Thomas-Nédélec mass matrix)	19
$B_{i,j}(\alpha)$	:=	$b(\mathbf{r}_i, s_j)$ (:= entry of discrete divergence operator in matrix form)	19
$\ker(A)$	:=	kernel of matrix or linear operator A	20
Z	:=	matrix with rows $\mathbf{z}_1, ..., \mathbf{z}_{\mathring{n}}$, where $\{\mathbf{z}_i\}$ is a basis of $\ker(B^T)$	20
$\mathring{A}$	:=	ZMZ^T	20
$\mathring{\mathbf{G}}$	:=	$Z\mathbf{G}$	20
$\mathring{\mathcal{V}}$	:=	$\{\mathbf{r}_h \in \mathcal{R}_h : \ b(\mathbf{r}_h, s_h) = 0 \text{ for all } s_h \in \mathcal{S}_h\}$	20

Chapter 3:			
X	:=	single random variable	25
$f(X)$	:=	probability density function	25
$\mathbb{E}[g(X)]$	:=	$\int_{-\infty}^{+\infty} g(x)f(x)\mathrm{d}x$ (:= expectation of function $g(X)$)	25
m	:=	$m_X := \mathbb{E}[X]$ (:= mean/expectation of X)	25
σ^2	:=	$\sigma_X^2 := \mathrm{Var}[X] := \mathbb{E}\left[(X-m)^2\right]$ (:= variance of X)	25
Y	:=	$\exp(Z)$ (:= lognormal Gaussian random variable for a Gaussian random variable Z)	26
$Z(\mathbf{x})$	:=	Gaussian random field for $\mathbf{x} \in \Omega$	26
$\mathbb{P}$	:=	probability	29
$\mathbf{Z}$	:=	$(Z(\mathbf{x}_1), ..., Z(\mathbf{x}_n))$ (:= vector of values of a random field Z at points $\mathbf{x}_1, ..., \mathbf{x}_n \in \Omega$)	29
C	:=	$\mathbb{E}\left[\mathbf{Z}^T\mathbf{Z}\right]$ (:= covariance matrix for random vector $\mathbf{Z}$)	30
L	:=	lower triangular matrix	30
$N_n(0, C(\mathbf{Z}))$	:=	vector of n iid. variables Z_i with mean 0 and and covariance $C(\mathbf{Z})$	32
$\#\mathcal{T}_\Omega^h$	:=	number of elements in $\mathcal{T}^h(\bar{\Omega})$	35
$\mathbf{x} + \mathcal{S}$	:=	$\{\mathbf{x}+\mathbf{y}: \quad \mathbf{y} \in \mathcal{S}, \quad \mathcal{S} \subseteq \mathbb{R}^d\}$	42
$\lvert\mathcal{S}\rvert$	:=	Lebesgue measure of a set $\mathcal{S} \in \mathbb{R}^d$	42
μ	:=	process rate	42
M_Ω	:=	$\sup_{\mathbf{x}\in\Omega}(Z(\mathbf{x}))$	42
m_Ω	:=	$\inf_{\mathbf{x}\in\Omega}(Z(\mathbf{x}))$	46

Symbol		Definition	Page
$\hat{M}_\Omega$	:=	$\inf_{\mathbf{x}\in\Omega}(\alpha(\mathbf{x}))$	46
$\hat{m}_\Omega$	:=	$\inf_{\mathbf{x}\in\Omega}(\alpha(\mathbf{x}))$	46
Chapter 4:			
α^τ	:=	$\frac{1}{\lvert\tau\rvert}\int_\tau \alpha$ (:= constant permeability value on τ)	53
$\text{supp}(\phi)$	:=	support of the function ϕ	53
$\mathcal{N}^h(\Omega)$	:=	set of degrees of freedom in Ω	53
$\lambda_{\max}(A), \lambda_{\min}(A)$	:=	maximum, resp. minimum eigenvalue of A	54
Chapter 5:			
$\mathcal{A}, \mathcal{A}_j$	:=	operator such that $\mathcal{A}: u \longmapsto -\text{div}(\alpha\nabla u(\mathbf{x}))$ for all $u \in C^2(\Omega)$	64
βH	:=	minimum overlap of the subdomains with $0 < \beta < 1$	66
$\mathcal{V}_j$	:=	$\left\{v_h \in \mathcal{V}^h : \text{supp}(v_h) \subset \widetilde{\Omega_j}\right\}$	67
$\mathcal{T}^H(\bar{\Omega})$	:=	coarse grid with elements Ω_i, $1 \le i \le p$ of size H	67
n_H	:=	cardinality of $\mathcal{T}^H(\bar{\Omega})$	67
$\mathcal{T}^h(\bar{\Omega}_i)$	:=	fine grid with elements τ of size h in Ω_i	67
n_j	:=	cardinality of $\mathcal{T}^h(\bar{\Omega}_j)$	67
R_j^T	:=	local extension operator, maps $\mathcal{V}_j \to \mathcal{V}^h$, $j = 1, ..., p$	67
A_j	:=	$R_j A R_j^T$, $j = 1, ..., p$, local stiffness matrix	67
$\hat{M}^{-1}$	:=	$R_1^T A_1^{-1} R_1 + R_2^T A_2^{-1} R_2$, one-level additive Schwarz preconditioner for two subdomains	69
M_1^{-1}	:=	$\sum_{j=1}^p R_j^T A_j^{-1} R_j$, one-level additive Schwarz preconditioner for several subdomains	69
$\bar{\Omega}_j$	:=	closure of an open subdomain $\Omega_j \subset \Omega$, $j = 1, ..., p$	69
ω_i	:=	interior $\left(\bigcup_{\Omega_j : \mathbf{x}_i^H \in \bar{\Omega}_j} \bar{\Omega}_j\right)$ with $\mathbf{x}_i^H \in \mathcal{N}^H(\Omega)$	70
ω_{Ω_j}	:=	$\bigcup_{k:\Omega_k \in \mathcal{T}^H(\bar{\Omega}), \Omega_k \cap \Omega_j \neq \emptyset} \Omega_k$ = neighbourhood of subdomains	70
$\mathcal{V}_0$	:=	space of coarse grid functions	70
$\{\psi_i^0\}$	:=	coarse grid basis functions	70
R_0^T	:=	$[\psi_1^0, ..., \psi_{n_H}^0]$, $n_h \times n_H$ global extension operator	70

Symbol		Definition	Page
A_0	:=	$R_0 A R_0^T$, coarse grid stiffness matrix	71
$M_{2,AS}^{-1}$	:=	$\sum_{j=0}^{p} R_j^T A_j^{-1} R_j = M_1^{-1} + R_0^T A_0^{-1} R_0$, two level add. Schwarz preconditioner	71
P_j	:=	$R_j A_j^{-1} R_j^T$, $j = 1, ..., p$, projections, self adjoint with respect to $a(.,.)$	71
$a(\mathbf{v}, \mathbf{w})$	:=	$\mathbf{w}^T A \mathbf{v}$	72
N_c	:=	minimum number of colours in the colouring argument	72
$\{\psi_i^{Lin}\}$	:=	linear coarse grid basis functions	77
$R_{0,Lin}^T$	:=	$[\psi_1^{Lin}, ..., \psi_{n_H}^{Lin}]$, $n_h \times n_H$ global linear extension operator	77
$A_{0,Lin}$	:=	$R_{0,Lin} A R_{0,Lin}^T$, linear coarse grid stiffness matrix	77
$M_{2,Lin}^{-1}$	:=	$M_1^{-1} + R_{0,Lin}^T A_{0,Lin}^{-1} R_{0,Lin}$, two level add. Schwarz preconditioner with lin. coarsening	77
$\{\chi_j\}$	:=	partition of unity, $j = 1, ..., p$	78
$B(d)$	:=	$\begin{cases} 1, & \text{if } d = 1 \\ (1 + \log(H/h)), & \text{if } d = 2 \\ (H/h), & \text{if } d = 3 \end{cases}$	80
I^h	:=	finite element interpolation onto $\mathcal{V}^h$	80
$\hat{\chi}_{i,\tau}$	:=	average of χ_i over τ	81
$\left(\bar{I}^H u\right)\left(\mathbf{x}_k^H\right)$	:=	quasi-interpolant $:= \begin{cases} 0, & \mathbf{x}_k^H \in \Gamma, \\ \lvert\omega_k\rvert^{-1} \int_{\omega_k} u(\mathbf{x}) \mathrm{d}x, & \text{otherwise}, \end{cases}$	82
Γ_j	:=	$\partial\Omega_j \backslash \Gamma$, $j = 1, ..., p$	86
$\hat{\Gamma}$	:=	$\bigcup_j \Gamma_j \backslash \Gamma$, $j = 1, ..., p$	86
$\Gamma_{\beta,j}$	:=	set of points in Ω_j within a distance βH of Γ_j, $j = 1, ..., p$	86

Chapter 6:			
$\{\psi_i^{Ms}\}$	:=	MsFE coarse grid basis functions	109
$R_{0,Ms}^T$	:=	$[\psi_1^{Ms}, ..., \psi_{n_H}^{Ms}]$, $n_h \times n_H$ global MsFE extension operator	109
$A_{0,Ms}$	:=	$R_{0,Ms} A R_{0,Ms}^T$, MsFE coarse grid stiffness matrix	109
$M_{2,Ms}^{-1}$	:=	$M_1^{-1} + R_{0,Ms}^T A_{0,Ms}^{-1} R_{0,Ms}$, two level additive Schwarz preconditioner with MsFE coarsening	109
$\mu_{i,j}$	:=	boundary condition of ψ_i^{Ms} on Ω_j	113

Symbol		Definition	Page
$E_{i,j}, E_{\mathbf{x},j}$	:=	coarse grid edge between and including $\mathbf{x}_i^H$ and $\mathbf{x}_j^H$, resp. $\mathbf{x}$ and $\mathbf{x}_j^H$	114
$\mathcal{H}_i$	:=	$\operatorname{diam}(\omega_i)$	118
$\gamma(\mathcal{V}_0, \alpha)$	:=	$\max_{\mathbf{x}_i^H \in \mathcal{N}^H(\Omega)} \left\{ \mathcal{H}_i^{2-d} \left\| \psi_i^0 \right\|_{H^1(\Omega),\alpha}^2 \right\}$ (:= coarse space robustness indicator)	118
n_S	:=	number of skeleton nodes in Ω	139
$R_{0,Skel}^T$	:=	$\left[\psi_1^{Ms}, ..., \psi_{n_s}^{Ms}\right]$, $n_h \times n_S$ global skeleton MsFE extension operator	140
$A_{0,Skel}$	:=	$R_{0,Skel} A R_{0,Skel}^T$, skeleton-based MsFE coarse grid stiffness matrix	140
$M_{2,Skel}^{-1}$	:=	$M_1^{-1} + R_{0,Skel}^T A_{0,Skel}^{-1} R_{0,Skel}$, two level additive Schwarz preconditioner with skeleton-based MsFE coarsening	140
$\mathcal{I}, \mathcal{S}$	:=	set of indices of fine grid nodes inside Ω, but not on the skeleton, resp. inside Ω on the skeleton	141

Chapter 7:			
n	:=	number of fine grid nodes in each coordinate direction	149
m	:=	number of coarse grid nodes in each coordinate direction such that n/m is an integer	149

Bibliography

[1] J. Aarnes and T. Y. Hou. *An Efficient Domain Decomposition Preconditioner for Multiscale Elliptic Problems with High Aspect Ratios.* Acta Mathematicae Applicatae Sinica, 18:63–76, 2002.

[2] R. Ababou. *Three-dimensional flow in random porous media.* PhD thesis, Mass. Instit. of Technol., 1988.

[3] R. Adler. *Excursions above high levels by Gaussian random fields.* Stoch. Proc. Appl., 5:21–25, 1977.

[4] R. J. Adler. *The Geometry of Random Fields.* Wiley Series In Probability And Mathematical Statistics, Salisbury, 1981.

[5] D. Aldous. *Probability Approximations via the Poisson Clumping Heuristic.* Springer-Verlag, New York, 1989.

[6] I. Babuška and J. E. Osborn. *Generalized Finite Element Methods: Their Performance and Their Relation to Mixed Methods.* SIAM J. Numer. Anal., 20:510–536, 1983.

[7] I. Babuška, G. Caloz, and J. E. Osborn. *Special finite element methods for a class of second order elliptic problems with rough coefficients.* SIAM J. Numer. Anal., 31:945–981, 1994.

[8] S. Balay, W. D. Gropp, L. C. McInnes, and B. F. Smith. *PETSc 2.0 User Manual.* Argonne National Laboratory, http://www.mcs.anl.gov/petsc/, 1997.

[9] S. Balay, K. Buschelman, W. D. Gropp, D. Kaushik, L. C. McInnes, and B. F. Smith. PETSc home page. http://www.mcs.anl.gov/petsc, 2001.

[10] G. Bennett. *Introduction to Ground-Water Hydraulics: A Programed Test for Self-Instruction.* U.S. Geol. Survey, Techniques of Water-Resources Investigations, Book 3, Chapter B2, 172pp., 1976.

[11] P. Billingsley. *Probability and measure, 3rd. edn.* Wiley, New York, 1995.

[12] P. E. Bjørstad and M. Skogen. Domain decomposition algorithms of Schwarz type, designed for massively parallel computers. In D. E. Keyes, T. F. Chan, G. A. Meurant, J. S. Scroggs, and R. G. Voigt, editors, *Fifth International Symposium on Domain Decomposition Methods for Partial Differential Equations*, pages 362–375, Philadelphia, PA, 1992. SIAM.

[13] P. E. Bjørstad, R. Moe, and M. Skogen. Parallel domain decomposition and iterative refinement algorithms. In W. Hackbusch, editor, *Parallel Algorithms for PDEs, Proceedings of the 6th GAMM-Seminar held in Kiel, Germany, January 19–21, 1990*, Braunschweig, Wiesbaden, 1990. Vieweg-Verlag.

[14] P. E. Bjørstad, M. Dryja, and E. Vainikko. Parallel implementation of a Schwarz domain decomposition algorithm. In J. Wasniewski, J. Dongarra, K. Madsen, and D. Olesen, editors, *Applied Parallel Computing in Industrial Problems and Optimization.* Springer, December 1996. URL `http://www.ii.uib.no/ petter/reports/para96.ps.gz`. Lecture Notes in Computer Science volume 1184.

[15] J. H. Bramble, J. E. Pasciak, and A. H. Schatz. *The construction of preconditioners for elliptic problems by substructuring, I.* Mathematics of Computations, 47:103–134, 1986.

[16] J. H. Bramble, J. E. Pasciak, J. Wang, and J. Xu. Convergence estimates for product iterative methods with applications to domain decomposition. *Math. Comp.*, 57(195):1–21, 1991.

[17] R. L. Bras and I. Rodriguez-Iturbe. *Random functions and hydrology.* Addison-Wesley, Reading, Massachussetts, 1985.

[18] L. Breiman. *Probability.* Addison-Wesley, MA, 1968.

[19] F. Brezzi and M. Fortin. *Mixed and Hybrid Finite Element Methods.* Springer, New York, 1991.

[20] P. Brockwell and R. Davis. *Time Series: Theory and Methods.* Springer, New York, 1987.

[21] P. Brooker. *Two-Dimensional Simulation by Turning Bands.* Mathematical Geology, 17:81–90, 1985.

[22] X.-C. Cai. *Some Domain Decomposition Algorithms for Nonselfadjoint Elliptic and Parabolic Partial Differential Equations.* PhD thesis, Mathematics Department, Courant Institute of Mathematical Sciences, New York, 1989.

[23] X.-C. Cai. An additive schwarz algorithm for nonselfadjoint elliptic equations. In T. Chan, R. Glowinski, J. Périaux, and O. Widlund, editors, *Third International Symposium on Domain Decomposition Methods for Partial Differential Equations*, pages 376–385, Philadelphia, PA, 1990. SIAM.

[24] X.-C. Cai, W. D. Gropp, and D. E. Keyes. A comparison of some domain decomposition and ILU preconditioned iterative methods for nonsymmetric elliptic problems. *Numer. Lin. Alg. Appl*, 1(5):477–504, 1994.

[25] A. Cangiani and E. Suli. Enhanced rfb method. Technical report, University of Oxford, 2003.

[26] G. Chan. An effective method for simulating gaussian random fields. *American Statistical Association 1999 Proceedings of the Statistical Computing Section*, pages 133–128, 1999.

[27] T. F. Chan. Parallel complexity of domain decomposition methods and optimal coarse grid sizes. *Parallel Computing*, 21(7):1033–1049, 1995.

[28] T. F. Chan and T. P. Mathew. *Domain Decomposition Survey.* Acta Numerica 1994, pages 61–143, 1994.

[29] H. Chen and T. Y. Hou. *A Mixed Multiscale Finit Element Method For Elliptic Problems with Oscillating Coefficients.* Journal of Mathematics of Computation, 72:541–576, 2002.

[30] G. Christakos. *Stochastic Simulation of Spatially Correlated Geoprocesses.* Mathematical Geology, 19:807–831, 1987.

[31] K. A. Cliffe, I. G. Graham, R. Scheichl, and L. Stals. *Parallel computation of flow in heterogeneous media using mixed finite elements.* Journal of Computational Physics, 164:258–282, 2000.

[32] N. Cressie. *Statistics for Spatial Data.* Wiley, New York, 1991.

[33] G. Dagan and S. Neumann. *Subsurface Flow and Transport: A Stochastic Approach.* Cambridge University Press, Cambridge, 1997.

[34] H. P. Darcy. *Les Fountaines Publiques de la Ville de Dijon.* Victon Dalmont, 1856.

[35] R. Davis and D. Harte. *Tests for Hurst Effect.* Biometrika, 74:95–101, 1987.

[36] A. Dembo, C. Mallows, and L. Shepp. *Embedding nonnegative definite Toeplitz matrices in nonnegative definite circulant matrices, with application to covariance estimation.* IEEE Transactions on Information Theory, 35:1206–1212, 1989.

[37] J. W. Demmel. *Applied Numerical Linear Algebra.* SIAM, Philadelphia, PA, 1997.

[38] C. Dietrich and G. Newsam. *A Fast and Exact Method for Multidimensional Gaussian Stochastic Simulations.* Water Resources Research, 29(8):2861–2869, 1993.

[39] M. Dorobantu and B. Engquist. *Wavelet-based numerical homogenization.* SIAM J. Numerical Analysis, 35:540–559, 1998.

[40] M. Dryja. Multilevel methods for elliptic problems with discontinuous coefficients in three dimensions. In D. E. Keyes and J. Xu, editors, *Seventh International Conference of Domain Decomposition Methods in Scientific and Engineering Computing*, volume 180 of *Contemporary Mathematics*, pages 43–47. AMS, 1994. Held at Penn State University, October 27-30, 1993.

[41] M. Dryja. A method of domain decomposition for 3-D finite element problems. In R. Glowinski, G. H. Golub, G. A. Meurant, and J. Périaux, editors, *First International Symposium on Domain Decomposition Methods for Partial Differential Equations*, pages 43–61, Philadelphia, PA, 1988. SIAM.

[42] M. Dryja and O. B. Widlund. *Some domain decomposition algorithms for elliptic problems.* Iterative Methods for Large Linear Systems, pages 273–291, 1989.

[43] M. Dryja and O. B. Widlund. *Towards a unified theory of domain decomposition algorithms for elliptic problems.* Third Int. Symp. on Domain Decomposition Methods for Partial Differential Equatios, 1990.

[44] M. Dryja and O. B. Widlund. *Domain decomposition algorithms with small overlap.* Technical Report 606, 1992.

[45] M. Dryja, M. V. Sarkis, and O. B. Widlund. Multilevel Schwarz methods for elliptic problems with discontinuous coefficients in three dimensions. *Numer. Math.*, 72(3):313–348, 1996.

[46] R. Dudley. *Gaussian processes on several parameters.* Ann. Math. Statist., 36: 771–788, 1965.

[47] R. M. Dudley. *Real Analysis and Probability.* Cambridge University Press, Cambridge, 2002.

[48] W. E. *Homogenization of linear and nonlinear transport equations.* Comm. Math. Sci., 1(1):87–133, 1992.

[49] W. E and B. Engquist. *The heterogeneous multi-scale method for homogenization problems.* Multiscale Modelling and Simulation, 2003.

[50] W. E and B. Engquist. *Multiscale modeling and computation.* Notices of AMS, 2003.

[51] W. E and B. Engquist. *The heterogeneous multi-scale method.* Comm. Math. Sci., 1(1):87–133, 2003.

[52] Y. R. Efendiev. *The Multiscale Finite Element Method (MsFEM) and Its Applications.* PhD thesis, California Inistitute of Technology, Pasadena, California, 1999.

[53] Y. R. Efendiev, L. J. Durlofsky, and S. H. Lee. *Modeling of Subgrid Effects in Coarse Scale Simulations of Transport in Heterogeneous Porous Media.* Preprint, 1999.

[54] Y. R. Efendiev, T. Y. Hou, and X.-H. Wu. *Convergence of a nonconforming multiscale finite element method.* SIAM Journal Numerical Analysis, 37(3):888–910, 2000.

[55] B. Engquist and O. Runborg. *Wavelet-based numerical homogenization with applications.* Multiscale and Multiresolution Methods: Theory and Applications, Lecture Notes in Computational Sciences and Engineering, 20:97–148, 2002.

[56] W. Feller. *An Introduction to Probability Theory and Its Applications.* John Wiley and Sons, Inc., New York, 1966.

[57] A. A. Freeze and J. A. Cherry. *Groundwater.* Prentice-Hall, New York, 1979.

[58] G. H. Golub and C. F. Van Loan. *Matrix Computations.* The Johns Hopkins University Press, Baltimore, Maryland, second edition, 1991.

[59] I. G. Graham and P. O. Lechner. Domain decomposition for heterogeneous media. In D. E. Keyes and O. Widlund, editors, *Sixteenth International Symposium on Domain Decomposition Methods for Partial Differential Equations (to be published)*, New York, 2006. Springer.

[60] I. G. Graham, P. O. Lechner, and R. Scheichl. *Domain Decomposition for Multiscale PDEs.* Preprint 11/06, Bath Institute For Complex Systems, 2006.

[61] D. Griffiths. *The construction of approximately divergence-free finite elements.* In *The Mathematics of Finite Elements and its Applications*, volume 3, pages 237–245, New York, 1979. Academic Press.

[62] G. Grimmett and D. Welsh. *Probability, an introduction.* Clarendon Press, Oxford, 1986.

[63] W. D. Gropp. Parallel computing and domain decomposition. In D. E. Keyes, T. F. Chan, G. A. Meurant, J. S. Scroggs, and R. G. Voigt, editors, *Fifth International Symposium on Domain Decomposition Methods for Partial Differential Equations*, pages 349–361, Philadelphia, PA, 1992. SIAM.

[64] W. D. Gropp and D. E. Keyes. A comparison of domain decomposition techniques for elliptic partial differential equations and their parallel implementation. *SIAM J. Sci. Stat. Comput.*, 8(2):s166–s202, 1987.

[65] W. D. Gropp and D. E. Keyes. Domain decomposition on parallel computers. *Impact Comput. Sci. Eng.*, 1:421–439, 1989.

[66] W. D. Gropp and D. E. Keyes. Parallel domain decomposition and the solution of nonlinear systems of equations. In R. Glowinski, Y. A. Kuznetsov, G. A. Meurant, J. Périaux, and O. Widlund, editors, *Fourth International Symposium on Domain Decomposition Methods for Partial Differential Equations*, pages 373–381, Philadelphia, PA, 1991. SIAM.

[67] P. Hall. *Introduction to the Theory of Coverage Processes.* Wiley, New York, 1988.

[68] T. Harter. *Unconditional and Conditional Simulation of Flow and Transport in Heterogeneous, Variably Saturated Porous Media.* PhD thesis, University of Arizona, 1994.

[69] J. Hawkes. *Local properties of some Gaussian processes.* Z. Wahrscheinlichkeitstheorie und Verw. Gebiete, 40:309–315, 1976.

[70] F. Hecht. *Construction d'une base d'un élement fini P_1 non conforme à divergence nulle dans $\mathbb{R}^3$.* PhD thesis, Université de Paris, 1980.

[71] F. Hecht. *Construction d'une base de fonctions P_1 non-conformes à divergence nulle dans $\mathbb{R}^3$.* RAIRO Modelisation Mathematique et Analyse Numérique, 15(2): 119–150, 1981.

[72] P. Hoel, S. Port, and C. Stone. *Introduction to probability theory.* Houghton Mifflin, Boston, 1971.

[73] T. Y. Hou. Numerical approximation to multiscale solutions in partial differential equations. In J. Blowey, A. Craig, and T. Shardlow, editors, *Frontier in Numerical Analysis*, pages 241–302, New York, 2003. Springer.

[74] T. Y. Hou. *Multiscale Modeling and Computation of Incompressible Flow.* In I. Sloan, editor, *Proceedings of the Fifth International Congress of Industrial and Applied Mathematics*, Philadelphia, 2003. SIAM.

[75] T. Y. Hou and X.-H. Wu. *A Multiscale Finite Element Method for Elliptic Problems in Composite Materials and Porous Media.* IMA Journal of Computational Physics, 134:169–189, 1997.

[76] T. Y. Hou, X.-H. Wu, and Z. Cai. *Convergence of a multiscale finite element method for elliptic problems with rapdily oscillating coefficients.* Journal of Mathematics of Computation, 68:913–943, 1999.

[77] M. K. Hubbert. *The Theory of Ground-Water Motion and Related Papers.* Hafner, New York, 1969.

[78] T. Hughes. *Multiscale phenomena: Green's functions, the Dirichlet-to-Neumann formulation, subgrid scale models, bubbles and the origins of stabilized methods.* Comp. Methods Appl. Mech. Engrg., 127:387–401, 1995.

[79] C. Johnson. *Numerical Solutions of Partial Differential Equations by the Finite Element Method.* Cambridge University Press, Cambridge, 1987.

[80] A. Jüngel and A. Unterreiter. *Discrete minimum and maximum principles for finite element approximations of non-monotone elliptic equations.* Numerische Mathematik, 99:485–508, 2005.

[81] M. Kac and D. Slepian. *Large excursions of Gaussian processes.* Ann. Math. Statist., 30:1215–1228, 1959.

[82] O. Kallenberg. *Foundations of modern probability.* Springer, Berlin, 1997.

[83] B. Kirk, J. Peterson, R. Stogner, and S. Petersen. *libmesh - a C++ finite element library, ver. 0.4.3.* Available at http://libmesh.sourceforge.net, 2004.

[84] B. Kozintsev. *GAUSSIAN - User manual.* University of Maryland, 1999.

[85] B. Kozintsev. *Computations With Gaussian Random Fields.* PhD thesis, University of Maryland, 1999.

[86] H. Kuchling. *Taschenbuch der Physik.* Fachbuchverlag Leipzig, Leipzig, 1991.

[87] M. R. Leadbetter, G. Lindgren, and H. Rootzen. Extremes and Related Properties of Random Sequences and Processes. *Springer Verlag*, New York, 1983.

[88] P. Lévy. *A special problem of Brownian motion and a general theory of Gaussian random functions.* Proc. Third Berkeley Symp. Math. Statist. Prob., 2:133–175, 1956.

[89] J. Liberty. *Teach Yourself C++ in 21 Days.* SAMS, Indianapolis, 2001.

[90] G. Lindgren. *Local maxima of Gaussian fields.* Ark. Math., 10:195–218, 1972.

[91] P.-L. Lions. *On the Schwarz alternating method I.* First International Symp. on Domain Decomposition Methods for Partial Differential Equations, 1988.

[92] P.-L. Lions. *On the Schwarz alternating method II.* Second International Symp. on Domain Decomposition Methods, 1989.

[93] P.-L. Lions. *On the Schwarz alternating method III: a variant for nonoverlapping subdomains.* Third International Symp. on Domain Decomposition Methods, 1990.

[94] A. Mack. *An element level zero-divergence finite element approach.* International Journal for Numerical Methods in Fluids, 19:795–813, 1994.

[95] A. Manatoglou. *The Digital Simulation of Multivariate Two- and Three-dimensional Stochastic Processes with a Spectral Turning Bands Method.* Mathematical Geology, 19(2):129–149, 1987.

[96] A. Manatoglou and J. Wilson. *The Turning Bands Method for Simulation of Random Field Using Line Generation by a Spectral Method.* Water Resources Research, 18(5):1379–1394, 1982.

[97] J. Mandel. Hybrid domain decomposition with unstructured subdomains. In A. Quarteroni, Y. A. Kuznetsov, J. Périaux, and O. B. Widlund, editors, *Domain Decomposition Methods in Science and Engineering: The Sixth International Conference on Domain Decomposition*, volume 157 of *Contemporary Mathematics*, pages 103–112. AMS, 1994. Held in Como, Italy, June 15–19, 1992.

[98] G. Matheron. *The Intrinsic Random Functions and Their Applications.* Advances in Applied Probability, 5:439–468, 1973.

[99] A. Matsokin and S. Nepomnyaschikh. *A Schwarz alternating method in a subspace.* Soviet Math., 29:78–84, 1985.

[100] J. W. Mercer and C. R. Faust. *Ground-Water Modelling: An Overview.* Ground Water, 18(2), 1980.

[101] J. W. Mercer and C. R. Faust. *Ground-Water Modelling: Mathematical Models.* Ground Water, 18(3), 1980.

[102] J. W. Mercer and C. R. Faust. *Ground-Water Modelling: Applications.* Ground Water, 18(5), 1980.

[103] J. Nečas. *Les méthodes directes en théorie des équations elliptiques.* Academia, Prague, 1967.

[104] B. Nédélec. *Mixed finite elements in* $\mathbb{R}^3$. Numerische Mathematik, 35:315–341, 1980.

[105] B. Nédélec. *Élements finis mixtes incompressibles pour l'équation de Stokes dans* $\mathbb{R}^3$. Numerische Mathematik, 39:97–112, 1982.

[106] S. Nepomnyaschikh. *Domain Decomposition and Schwarz Methods in a Subspace for the Approximate Solution of Elliptic Boundary Value Problems.* PhD thesis, Computing Center of the Siberian Branch of the USSR Academy of Sciences, Novosibirsk, USSR, 1986.

[107] Numerical Analysis Group Limited. *NAG Fortran Library Manual Mark 14.* Oxford, 1990.

[108] J. Pickands. *Asymptotic properties of the maximum in a stationary Gaussian process.* Trans. Amer. Math. Soc., 145:75–86, 1969.

[109] T. Prickett. *Modeling Techniques for Groundwater Evaluation.* Advances in Hydroscience, 10:1–143, 1975.

[110] C. Qualls and H. Watanabe. *Asymptotic properties of Gaussian random fields.* Trans. Amer. Math. Soc., 177:155–171, 1973.

[111] P. A. Raviart and J. M. Thomas. A mixed finite element method for 2-nd order elliptic problems. In A. Dold and B. Eckmann, editors, *Mathematical Aspects of Finite Element Methods.* Springer, 1977. Lecture Notes of Mathematics, Volume 606.

[112] B. Ripley. *Stochastic Simulation.* Wiley, New York, 1987.

[113] S. Ross. *A first course in probability.* Prentice-Hall, New York, 1998.

[114] Y. Rozanov. *On Gaussian fields with given conditional distributions.* Theor. Probability Appl., 12:381–391, 1967.

[115] G. Sangalli. *Capturing small scales in elliptic problems using a Residual-Free Bubbles finite element method.* Multiscale Modeling and Simulation, 1:485–503, 2003.

[116] R. Scheichl. *Iterative Solution of Saddle Point Problems Using Divergence-free Finite Elements with Applications to Groundwater Flow.* PhD thesis, University of Bath, 2000.

[117] R. Scheichl. Decoupling three-dimensional mixed problems using divergence-free finite elements. *SISC*, 23, 2002.

[118] H. A. Schwarz. Über einen Grenzübergang durch alternierendes Verfahren. *Vierteljahrsschrift der Naturforschenden Gesellschaft in Zürich*, 15:272–286, May 1870.

[119] A. Shiryayev. *Probability.* Springer, Berlin, 1984.

[120] M. Skogen. *Schwarz Methods and Parallelism.* PhD thesis, Department of Informatics, University of Bergen, Norway, 1992.

[121] B. F. Smith. *Domain decomposition algorithms for the partial differential equations of linear elasticity.* PhD thesis, Courant Institute of Mathematical Sciences, Department of Computer Sciences, 1990.

[122] B. F. Smith. An optimal domain decomposition preconditioner for the finite element solution of linear elasticity problems. *SIAM J. Sci. Comp.*, 14(2):361–378, 1992.

[123] S. L. Sobolev. *L'Algorithme de Schwarz dans la Theorie de l'Elasticite.* Comptes Rendus (Doklady) de l'Academie des Sciences de l'URSS, (IV(XIII)6):243–246, 1936.

[124] D. Stirzaker. *Probability and random variables.* Cambridge University Press, Cambridge, 1999.

[125] F. Thomasset. Numerical solution of the navier-stokes equations by finite element methods. In *Von Karman Institute Lecture notes, no. 86 (Computational Fluid Dynamics, March 21-25, 1977)*, New York, 1977. Springer.

[126] F. Thomasset. *Implementation of Finite Element Methods for Navier-Stokes Equations.* Springer, New York, 1981.

[127] A. Tompson, R. Ababou, and L. Gelhar. *Implementation of the three dimensional Turning Bands random field generator.* Water Resources Research, 25(10):2227–2243, 1989.

[128] A. Toselli and O. B. Widlund. *Domain Decomposition Methods - Algorithms and Theory.* Springer, New York, 2005.

[129] E. Vanmarcke. *Random Fields: Analysis and Synthesis.* The MIT Press, London, 1983.

[130] W. Venables, D. Smith, and the R Development Core Team. *An Introduction to R - Notes on R: A Programming Environment for Data Analysis and Graphics.* R Development Core Team, 2003.

[131] H. F. Wang and M. P. Anderson. *Introduction to Groundwater Modelling - Finite Difference and Finite Element Methods.* W.H. Freeman and Company, San Francisco, 1982.

[132] A. Wood and G. Chan. *Simultation of stationary Gaussian Processes in* $[0,1]^d$. Journal of Computational and Graphical Statistics, 3(4):409–432, 1994.

[133] X. Ye and G. Anderson. *The derivation of minimal support basis functions for the discrete divergence operator.* Journal of Computational and Applied Mathematics, 61:105–116, 1995.